SEPARATING MORAL HAZARD FROM ADVERSE SELECTION AND LEARNING IN AUTOMOBILE INSURANCE: LONGITUDINAL EVIDENCE FROM FRANCE

Georges Dionne  
HEC Montreal  
Maki Dahchour  
Aviva Canada

Pierre-Carl Michaud  
Université du Québec à Montréal and RAND Corporation

Abstract
The identification of information problems in different markets is a challenging issue in the economic literature. In this paper, we study the identification of moral hazard from adverse selection and learning about risk within the context of a multi-period dynamic model. We extend the model of Abbring et al. (2003) to include learning about risk and insurance coverage choice over time. We derive testable empirical implications for panel data. We then perform tests using longitudinal data from France during the period 1995-1997. We find evidence of moral hazard among a sub-group of policyholders with less driving experience (less than 15 years). Policyholders with fewer than 5 years of experience have a combination of learning about risk and moral hazard, whereas no residual information problem is found for policyholders with more than 15 years of experience. (JEL: C33, C35, D82, D86, G22)

1. Introduction

Information asymmetries generally follow two distinct pathways. Adverse selection strongly predicts a positive correlation between the accident probability of a policyholder and the generosity of his insurance contract. 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.

The evidence is not conclusive concerning the existence of residual asymmetric information in automobile insurance markets. Some studies using the conditional correlation approach on cross-sectional data find evidence of asymmetric information (Puelz and Snow, 1994; Cohen, 2005) while others did not (Chiappori and Salanie, 2000; Dionne et al., 2001). 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 (Chiappori, 2000).²

Abbring et al. (2003) investigate the dynamics in claims as a way of directly testing for moral hazard. Under most experience rating regimes, an at-fault claim raises the cost of future claims. Hence, such regimes should promote safe driving, at least in theory. Empirically, negative occurrence dependence in accidents within a risk class should be observed under moral hazard. The authors find little evidence of moral hazard in France. In fact, there is only weak evidence among inexperienced drivers, which points to learning about risk rather than moral hazard, i.e. beginners who learn they are bad risks exert caution.³ To separate learning about risk leading to adverse selection (asymmetric learning) from moral hazard, we consider the case where information on contracts and accidents is available for multiple years in the form of panel data. We exploit dynamics in accidents and insurance coverage controlling for dynamic selection due to unobserved heterogeneity. Changes in insurance coverage allow us to construct two additional tests, one for moral hazard and the other for asymmetric learning. Coupled with the negative occurrence test of Abbring et al. (2003), this trilogy of tests allows us to separate moral hazard from asymmetric learning.

We analyze the identification of asymmetric learning and moral hazard within the context of a tractable structural dynamic insurance model. From the solution of the model, we simulate a panel of drivers behaving under different information regimes or data generating processes. We validate our empirical tests on simulated data generated from these different information regimes. We then apply these tests to longitudinal data on accidents, contract choice and experience rating for the period 1995-1997 in France. We find no evidence of information problems among experienced drivers (more than 15 years of experience). For drivers with less than 15 years of experience, we find strong evidence of moral hazard but little evidence of asymmetric learning. We obtain evidence of asymmetric learning, despite the small sample size, when focusing on drivers with less than 5 years of experience.

2. Other results in insurance markets are summarized in Cohen and Siegelman (2010).
3. Also, with cross-sectional data, unobserved confounders such as risk aversion may mask evidence of asymmetric information (Finkelstein and McGarry, 2006) or advantageous selection (Fang, Keane and Silverman, 2008). Chiappori et al. (2006) propose cross-sectional tests based on profit maximization in competitive markets that are robust to differences in risk aversion.
4. Their estimate of negative occurrence dependence among inexperienced drivers is insignificant. Further, combining claims at-fault and other claims changes the result significantly, as would be predicted under asymmetric learning (all accidents should matter). Dionne et al. (2011) extend their test based on demerit points in Canada and find evidence of moral hazard. Abbring et al. (2008) analyze accidents and claims reporting in a dynamic setting with moral hazard using data from the Netherlands. They also find evidence of moral hazard.
The remainder of the paper is structured as follows. In Section 2, we present the theoretical model we use to construct empirical tests. In Section 3, we present our empirical tests and validate them using simulated data from the theoretical model in Section 2. In Section 4, we present results of the tests applied to French panel data. Section 5 concludes. An online appendix provides additional information on the data and results.

2. Theoretical Model

To investigate how contract choices and accident outcomes allow us to distinguish moral hazard and asymmetric learning leading to adverse selection, we start with a dynamic model of moral hazard. We build a model where policyholders directly choose the probability of future accidents by exerting effort. In order to consider other information problems in a multi-period context, we extend the model along two dimensions.

First, we add a contract choice decision to the model. Second, we consider the possibility that there is learning about risk (drivers and insurers learn about an individual’s innate risk). Over time, this may lead to adverse selection if the policyholder learns faster than the insurer, or to full information if both share symmetric learning. Because the data we will use contains all accidents (not only claims), we allow consumers to learn faster than insurers, hence allowing for asymmetric learning. A final characteristic is that we use a discrete-time model with one period lasting one year. A higher-frequency model where effort can be adjusted continuously within the contract period would require information on contract renewal and accident dates which we do not have in the data. But, a continuous time model could be used to simulate discrete-time data.

Adding these extensions to the pure moral hazard model makes analytical solutions difficult, so we solve the model numerically and show how the dynamics in accidents and insurance coverage allow us to separate moral hazard from adverse selection and learning about risk. Because the model has clear policy parameters that govern the presence of moral hazard, adverse selection and learning about risk, we can simulate a cohort of drivers from different scenarios and confirm whether the empirical tests we propose can separate these information problems.

The most important insurance decision in France is that of buying comprehensive insurance coverage (CC) in addition to compulsory “responsabilité civile” or limited liability coverage (LL). By law, every driver must purchase an LL contract that protects third parties if the driver is at fault. However, the LL contract does not cover the driver’s damages if he is at fault. The CC contract covers such damages and the policyholder pays a deductible only when he is at fault, which varies across contracts. Finally, the other party’s insurer pays all damages if the policyholder is not at fault. The insurer observes only the claims, whether the insured is at fault or not, while the insured observes all accidents.

The total premium paid is experience rated. It is scaled by a coefficient, the bonus-malus coefficient, which is a function of the history of claims where the
driver is at fault. We assume drivers differ in terms of risk type (or ability). In the model, agents first buy insurance without knowledge of their own risk. They learn about risk from their history of accidents. Accidents differ depending on whether the driver is at fault or not. Although the insurer observes the bonus-malus he does not learn as fast as the agent about his riskiness, partly because he/she observes claims only. Thus asymmetric learning develops, which may lead to pure adverse selection in contract choices within observable risk classes (Rothschild and Stiglitz, 1976; Crocker and Snow, 1986). The agent can influence his accident probability by exerting effort. Effort is unobservable to the insurer and there is moral hazard within a given risk class because the driver has less incentive to exert effort under more generous insurance contracts (Holmstrom, 1979; Shavell, 1979) or when past accidents increase the cost of future claims (Abbring et al., 2003).

Contracts are renewed annually. An agent makes two simultaneous decisions at each period \( t = 1, ..., T \) consisting of the choice of contract and of the level of effort to prevent accidents. We do not allow these choices to be sequential. This would be more relevant in a higher-frequency model where effort can be adjusted continuously within the contract period. Time in the model represents driving experience. The timing of the decisions in the model is represented in Figure 1.

The agent pays the premium on a contract when he makes his contract decision, conditional on past accidents and beliefs. After the agent has made his decisions, uncertainty is resolved. We assume only one accident can occur in each period (few drivers have more than one accident per year in the data). We denote by \( n_t = \{0, 1, 2\} \) the occurrence of accidents where 0 means no accidents, 1 denotes an accident where the driver is not at fault and 2 an accident where he is at fault. As usual \( n_t \) is unknown prior to making decisions at \( t \). At the beginning of the following period damages are paid prior to the agent’s making any new decision.

We denote the choice of the CC contract as \( d_t^c = 1 \); \( d_t^c = 0 \) when only the LL contract is retained as coverage. An accident results in a fixed monetary loss, \( L \). If \( d_t^c = 1 \) and if the policyholder has an at-fault accident, he pays, at most, the deductible \( f < L \). If he does not purchase CC, he pays damage \( L \). If the driver is not at fault, no payments are made by the driver, regardless of whether CC is purchased.

---

5. The French experience rating system rates at-fault claims (Dionne, 2001). The information on the bonus-malus is public and shared across insurers. The system is enforced by a law stipulating the penalty (malus) in case of an at-fault accident and reward (bonus) otherwise. The rating coefficient, or the bonus-malus coefficient, is applied proportionally to the base premium at the time the signing/renewal of the contract.
FIGURE 1. Timing of the model and contract. The figure shows the timing assumed in the model. Model periods and contract periods do not coincide. In the model, individuals start by observing whether they had an accident or not; damages are paid by the insurer if an accident occurred. This is for a pre-determined contract which they have chosen in the previous year. Based on the occurrence of the accident but before renewing their contract, they update their beliefs on whether they are high risk (equation 5 in text) or not and the insurer updates the bonus-malus (equation 1 in text). Uncertainty about income is resolved before they choose a new contract (at the same time accidents occur). Based on the information they have at that point, they choose to renew or change their insurance contract. They also choose the effort level for the following year. This is the start of the contract period $t$. The model period $t$ ends when an accident based on the effort level and insurance contract chosen in $t$ occurs. In $t+1$, a new period model starts with the damages paid depending on the insurance contract and the occurrence of an accident. Beliefs are then revised, the bonus-malus updated and the individual renews the contract.

The premium has two components: an *a priori* and an *a posteriori* pricing component. The *a posteriori* component is a function of the driver’s accident history, summarized by his bonus-malus $b_{t-1}$ in preceding contract period, and the occurrence of an at-fault accident during the most recent contract period. The new bonus-malus for the current contractual year is updated according to

$$b_t = b(b_{t-1}, n_{t-1}) = \begin{cases} \delta_b b_{t-1} & \text{if } n_{t-1} \neq 2 \\ \delta_m b_{t-1} & \text{if } n_{t-1} = 2 \end{cases}$$

(1)

where $\delta_b$ is the bonus coefficient (0.95) and $\delta_m$ is the malus coefficient (1.25). In the French market, the first at-fault accident does not increase the premium if the policyholder has the minimum coefficient (0.5). We keep track of this clause in the model.

The *a priori* pricing component depends on the choice of the CC contract. The total premium paid ($pr$) is the product of the bonus-malus and the *a priori* component

$$pr(b_t, d_t) = b_1 \exp(\rho_0 + \rho_1 d_t)$$

(2)

where $\exp(\rho_0) > 0$ denotes the percentage increase in the premium for the CC contract and $\exp(\rho_0)$ is the base premium for the LL coverage.
We assume it is costly to change coverage over time. The cost is given by \( \psi \) and is symmetric (the same for increasing and decreasing coverage). Contract choices are quite persistent in the data and not allowing for switching costs would entail too many transitions compared with the data. Furthermore, some studies show that price dispersion is large for the same insurance product across insurers, which is consistent with the existence of switching or search costs (see Schlesinger and von der Shulenburg, 1993).

2.1. Effort and Accidents

Agents can choose to exert prevention effort \( e_t = \{0, 1\} \) during the contractual year. Effort reduces the probability of both types of accident. Assume the probability of an accident of type \( j \) takes the multinomial logit form

\[
p(n_t = j | e_t, \alpha) = \frac{\exp(\mu_{j0} + \mu_{j1}e_t + \mu_{j2}\alpha)}{1 + \sum_{j=1,2} \exp(\mu_{j0} + \mu_{j1}e_t + \mu_{j2}\alpha)}, \quad j = 1, 2
\]  

where \( \mu_{j1} \) is the coefficient affecting the marginal productivity of effort for the probability of an accident of type \( j \) (no accident is the reference). The parameter \( \mu_{j2} \) is the factor loading for risk classification \( \alpha \). The individual is uncertain of his true innate risk classification \( \alpha \), which is time invariant. It can take two values \( \alpha = \{0, \alpha_H\} \), where \( \alpha_H \) denotes the high-risk type. The accident probabilities are serially uncorrelated over time, for a given \( \alpha \) and effort sequence. We also assume for simplicity that there is no “experience” effect irrespective of the accident history, i.e. accident probabilities do not depend on experience and thus there is no learning-by-doing in the model. However, there is learning about risk.

The experience of the policyholder can be used to construct expectations about his innate riskiness. He/she knows the form of equation (3) and its parameter values. The fraction of high types \( \alpha_H \) in the population is denoted \( \delta_H \). We denote by \( \pi_t(\alpha_H | n_{t-1}, e_{t-1}, e_{t-2}, \ldots, e_1, \alpha_H) \) the subjective probability that the policyholder is a high-risk type for the current contractual period given his effort and accident realization in the previous periods. We denote the accident history by \( n' = (n_1, \ldots, n_t) \) and do the same for other variables. Using Bayes’ rule, the probability that the driver is a high-risk type given his history or experience up to \( t \) is given by

\[
\Pr(\alpha_H | n^{t-1}, e^{t-1}) = \frac{\Pr(n_{t-1} | e_{t-1}, \alpha_H) \Pr(\alpha_H | n^{t-2}, e^{t-2})}{\Pr(n_{t-1} | e_{t-1})}
\]  

using equation (3). Thus, denoting \( \pi_t \) as the subjective belief in period \( t \), equation (4) simplifies to a recursive form
\[ \pi_t = \frac{\pi_{t+1} \Pr(n_{t+1} \mid e_{t+1}, \alpha)}{\pi_{t+1} \Pr(n_{t+1} \mid e_{t+1}, \alpha) + (1 - \pi_{t+1}) \Pr(n_{t+1} \mid e_{t+1}, 0)}. \]  

This equation shows how policyholders update their prior probability of being a particular risky type from \( \pi_{t+1} \) to \( \pi_t \). Having an accident makes it more likely they are the high-risk type. It makes it even more likely if they exerted effort but still had an accident.

At every period, the subjective probability of an accident of type \( j \) is given by
\[ \tilde{p}(n_t = j \mid e_t, \pi_t) = \pi_t p(n_t = j \mid e_t, \alpha) + (1 - \pi_t) p(n_t = j \mid e_t, 0), \quad j = 1, 2 \]  

2.2. Maximization Problem

Within a period, utility is assumed separable in consumption \( (c_t) \) and effort
\[ u(c_t, e_t) = \frac{c_t^{1-\sigma}}{1-\sigma} + \gamma e_t \]  

The coefficient of relative risk aversion is \( \sigma \) and the marginal disutility of effort is \( \gamma < 0 \).

Agents receive income \( y_t \) every period. We assume income is stochastic and follows a first order Markov process. The distribution \( F(y_{t+1} \mid y_t) \) is known to the agent. We also assume there are no assets in the model so that
\[ c_t = y_t - pr(d_t, h_t) - \psi I(d_t \neq d_{t-1}) - n_{2,t-1} L(d_{t-1}) \]  

where \( L(d_{t-1}) = (1 - d_{t-1}) L + d_{t-1} f \) and \( n_{2,t-1} \) is equal to 1 if an at-fault accident occurs and zero otherwise. The indicator function is given by \( I(z) = 1 \) if \( z \) is true and zero if not. Finally, we denote the agent’s subjective discount factor as \( \beta \).

For notational convenience, we partition the state space at \( t \) into the vectors \( s_t = (\pi_{t-1}, h_{t-1}, d_{t-1}, e_{t-1}) \) and \( (n_{t-1}, y_t) \). This partition is helpful when writing down the Bellman equation since at time \( t \), past accidents and current income are known but future realizations are not. Before decisions are made, the agent updates his beliefs using equation (5) and the insurer updates his bonus-malus using equation (1). Therefore, the final state-variables on which the agent relies to make his decisions are \( (\pi_t, h_t, d_{t-1}, y_t) \). Losses from accidents during the preceding contractual period are deducted from current income. After decisions are made, uncertainty about accidents and income for the next period is resolved.

The agent’s optimization problem can be expressed as a series of one-period problems using Bellman’s principle of optimality. Both effort and contract choice decisions at each time period are made simultaneously given current information. We have
\[ V_t(s_t, n_{t-1}, y_t) = \max_{e_t, d_t} u(c_t, e_t) + \beta \sum_{n} V_{t+1}(s_{t+1}, n, y_{t+1}) p(n_t = n \mid e_t, \pi_t) dF(y_{t+1} \mid y_t) \]  

for \( t = 1, \ldots, T \). The maximization problem can be solved by backward recursion from the last period \( T = 40 \). We discretize the state-space for the log of the bonus-malus \( b \).
over the interval \([0.5, 2]\) and \(\pi\) over the unit interval.\(^5\) We use 35 grid points for both. It is important to stress that the optimal solution for a given period will be a function of all variables at the beginning of the period. Hence, we will focus on the dynamics in contract choice and accidents to derive empirical tests of moral hazard and asymmetric learning.

2.3. Calibration

2.3.1. SOFRES Longitudinal Survey

The main data source we use is the SOFRES longitudinal survey, *Parc Automobile*, which is a rotating panel representative of the French policyholders and their vehicles for the years 1995, 1996 and 1997. Respondents were interviewed by mail (questionnaires were sent each January) about several topics including car insurance and their accident history. SOFRES is an independent survey organization conducting consumer surveys and is not an insurer. This allowed information gathering on both claims and accidents, including those not reported to insurers. Furthermore, SOFRES re-interviews individuals even if they switch insurers. Finally, the timing of the survey does not correspond exactly to the renewal of insurance contracts. Contracts are renewed year round in France while the survey is conducted each January and accidents over the previous 12 months are reported. Unfortunately, no information on contract renewal dates and dates for accidents is available in the dataset. Furthermore, we cannot distinguish at-fault claims from other claims. We discuss the problems associated with these features of the data in Section 3.2.

We define an observation as a respondent-vehicle pair. A sizeable proportion of contracts are observed for less than three years. SOFRES aims to maintain the representativity of its sample over time, which means that some respondents are not re-interviewed for exogenous reasons or reasons related to observable characteristics that we control for (i.e. region). We keep entries from 1996 and 1997 even if they were not present in 1995. The online appendix provides more details on the number of observations and the participation patterns in the survey.

The survey contains four modules. The first covers the socio-economic characteristics of policyholders. The second covers characteristics of the vehicle. In the online appendix we provide descriptive statistics on these data (see also Figure A.1). The third concerns insurance contracts. It provides the current bonus-malus coefficient and the current type of insurance coverage: CC or LL contract. The bonus-malus coefficient is updated at the end of every contractual year and is constant within a year (i.e. premiums can be adjusted only at the time of contract renewal). The new bonus-malus appears only when it is time to negotiate a new contract. Given the design of survey, this implies that the bonus-malus in a given

\(^6\) In theory, the bonus-malus coefficient is allowed to reach a maximum of 3.5. Very few observations in the SOFRES panel have a bonus-malus greater than 2 (4 observations out of 12,000). To save on computations, we therefore set the interval between 0.5 and 2.
year does not take account of accidents that have taken place during the months preceding the survey date but after the contract renewal date.

2.3.2. Calibration

We chose a baseline calibration such that both moral hazard and asymmetric learning are present. We do not claim to be able to estimate all parameters, such as risk aversion, from the data. Estimation given the data we have is a fairly ambitious endeavor and has been left for further research. We use parameter values that yield reasonable profiles comparable to those in the data and complement the SOFRES data with market data from the French Federation of Insurers (FFSA). Table 1 shows the parameters we use in the baseline scenario. All monetary amounts are in 2009 euros (thousands).

TABLE 1. Parameter values.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Definition</th>
<th>Value</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>$s$</td>
<td>coefficient of relative risk aversion</td>
<td>1.5</td>
<td>literature</td>
</tr>
<tr>
<td>$g$</td>
<td>marginal disutility effort</td>
<td>-5.00E-04</td>
<td>assumed</td>
</tr>
<tr>
<td>$m10$</td>
<td>constant, accident probability</td>
<td>-2.1</td>
<td>calibrated</td>
</tr>
<tr>
<td>$m11$</td>
<td>productivity effort, accident probability</td>
<td>-0.5</td>
<td>assumed</td>
</tr>
<tr>
<td>$m12$</td>
<td>factor high risk type, accident probability</td>
<td>0.75</td>
<td>assumed</td>
</tr>
<tr>
<td>$m20$</td>
<td>constant, at fault probability</td>
<td>-3.2</td>
<td>calibrated</td>
</tr>
<tr>
<td>$m21$</td>
<td>productivity effort, at fault probability</td>
<td>-0.5</td>
<td>assumed</td>
</tr>
<tr>
<td>$m22$</td>
<td>factor high risk, at fault probability</td>
<td>1.6</td>
<td>assumed</td>
</tr>
<tr>
<td>$f$</td>
<td>deductible</td>
<td>0.192</td>
<td>FFSA</td>
</tr>
<tr>
<td>$L$</td>
<td>loss</td>
<td>2.439</td>
<td>FFSA</td>
</tr>
<tr>
<td>$r0$</td>
<td>log premium LR</td>
<td>-1.69</td>
<td>FFSA</td>
</tr>
<tr>
<td>$r1$</td>
<td>log % change in premium for CC contract</td>
<td>0.7</td>
<td>FFSA</td>
</tr>
<tr>
<td>$b$</td>
<td>discount factor</td>
<td>0.985</td>
<td>assumed</td>
</tr>
<tr>
<td>$db$</td>
<td>bonus factor</td>
<td>0.95</td>
<td>FFSA</td>
</tr>
<tr>
<td>$dm$</td>
<td>malus factor</td>
<td>1.25</td>
<td>FFSA</td>
</tr>
<tr>
<td>$y$</td>
<td>switching or search cost (euros)</td>
<td>0.05</td>
<td>assumed</td>
</tr>
<tr>
<td>$d$</td>
<td>fraction of population high risk</td>
<td>0.3</td>
<td>assumed</td>
</tr>
</tbody>
</table>

Notes: This table presents parameters of the models along with the value used in the baseline simulations. The source for the value is also reported. Calibrated means that the parameter was chosen to match certain features of the data as defined in text. Assumed implies that values were chosen arbitrarily. FFSA stands for the "Federation Francaise des Societes d'Assurances".

Preference Parameters: We use a coefficient of relative risk aversion equal to 1.5. We calibrate this parameter so that it matches the fraction of policyholders with comprehensive coverage in the SOFRES panel. The value of 1.5 is at the low-end of the values reported in Attanasio and Weber (1995) [1.49-3.39]. Because the average loss and premiums are small relative to income in our application, a value of 1.5 is likely reasonable for our purposes. The marginal
disutility cost of effort is difficult to calibrate. We have assumed a value of -0.0005. The discount factor was assumed to be 0.985.

**Accident Process**: We calibrate the intercepts of the accident process such that it matches the fraction of accidents in the data. However, we do not observe at-fault claims in the SOFRES panel, only total claims, which include accidents where another party is at fault. Abbring et al. (2003) report that the annual rate of at-fault claims is 6.4% in their sample, which is also from France. We use this figure to calibrate the intercepts. We assume the effort coefficient is -0.5 for both types of accident probabilities (at-fault or not). This yields a sizeable effect on the probability of having an accident. Finally, we assume arbitrarily that the factor coefficient of the high risk-type \( \mu_{12} \) is 0.75 in the accident probability equation. The parameter \( \mu_{22} \) is set to 1.6 in the at-fault accident probability equation. Hence, we assume that high-risk types are more likely to have at-fault accidents than low-risk types. We assume, quite arbitrarily, that 30% of the population is high-risk.

**Market Data**: The SOFRES data set contains little information on premiums and deductibles, so we use aggregate information obtained from the FFSA for the year 1997. The average deductible \( (f) \) in France was quite low, 192 euros. The average loss \( (L) \) was 2,439 euros. The average premium for the LL coverage was 184 euros. The coefficient \( \rho_1 \) is set at 0.7 for the CC contract. The CC premium is 370 euros, which is double that of the LL premium. The bonus factor is 0.95 and the malus factor 1.25. We assume the switching cost is 50 euros.

**Income Process**: The SOFRES panel contains information on income but the information is categorical, and the bins are not necessarily natural ones. Instead, we use data from the French subsample of the European Community Household Panel (ECHP) for the years 1995-1997. We use household net income along with OECD equivalence scales to adjust for household composition. We discretize income into 10 categories (deciles) and use the midpoint within each category as the approximate value in the simulations. We then compute the transition matrix across these deciles to obtain an estimate of \( F(y_{t+1} | y_t) \). We use ECHP-provided weights when computing this matrix. The resulting median net (equivalized) income is 14,319 euros and there is a fair amount of persistence in income over time.

### 2.3.3. Simulation of the Baseline Scenario

We solve the model for the optimal decision rules given this choice of parameters. We use \( T=40 \) as the maximum experience level. We then simulate 2000 drivers until they reach 25 years of experience. When simulating individuals, we randomly draw the initial bonus-malus, subjective probability of being high risk, risk type and income. This constitutes the heterogeneity in the simulation. We draw the bonus-malus from the SOFRES empirical distribution for drivers with less than 2 years of experience. The average bonus-malus is roughly 0.75. The initial subjective

---

7. The OECD equivalence scale used in the ECHP is the sum of weighted household members where the weight is one for the first adult member, 0.5 for subsequent adults (over 18 years old) and 0.3 for each child (under 18 years old).
probability is drawn from a normal distribution. Let \( \Phi(a + c) \) be the standard normal cumulative distribution where \( c \) is a driver-specific draw from the standard normal distribution and \( a \) is a constant such that the average subjective probability is 0.3. This is a natural average prior given that 30% of the population is high risk. The risk type is drawn independently of all other variables. Since we know each driver’s type, either low- or high-risk, we compute statistics by risk type and experience.

We report the results of the simulations in Figure 2. As one would expect, there is a clear distinction between the behavior of high and low risk drivers. Differences grow larger with experience as both types learn what type of driver they are. This can be seen from the first panel reporting the subjective probability of being high type. As the high risk individuals have more accidents, they slowly become more convinced that they are high risk. The opposite occurs for low-risk types. Since the low risk types have less accidents, particularly at-fault accidents, their bonus-malus coefficient falls more rapidly towards 0.5. The bonus-malus coefficient of the high risk types remains high. In our model, the bonus-malus coefficient is an excellent indicator of risk.

The portion with CC coverage is initially very similar between the two groups. This is because both risk types do not know yet what kind of driver they are likely to be. As they become more convinced, the high-risk types purchase the CC coverage more rapidly than those who learn they are likely low-risk. Over time, a positive correlation emerges between risk and coverage, which leads to adverse selection, as predicted by asymmetric learning. A substantial proportion of the population exerts effort, and high risk drivers eventually make a greater effort to reduce their probability of having an accident. In terms of the joint distribution of effort and coverage, we find that very few drivers choose not to make an effort and forgo CC coverage (less than 5%). Most drivers choose either one or both options.

3. Empirical Tests

3.1. Optimal Decisions

To derive tests for moral hazard and asymmetric learning, we will focus on the dependence of decision rules on the state-space variables. The optimal effort level is a discrete function \( e_t^*(s_t, n_{t-1}) \). It depends on the state \( s_t = (\pi_{t-1}, b_{t-1}, d_{t-1}, e_{t-1}) \) and the realization of an accident in the previous period. Within the context of the model, moral hazard can be detected if the accident process depends on any of the state variables of the model. In the absence of effort, the accident process (equation 3) does not depend on state variables of the model. In particular, moral hazard is present if effort depends on contract parameters which leads to a dependence of accidents on contract parameters. The two contract

---

8. Effort is also a function of income. For notational convenience in this section we omit the dependence of effort on income.
parameters that vary between individuals are coverage choice and the bonus-malus coefficient that affects the premium.
Abbring et al. (2003) show within the context of the model that optimal effort increases with the premium via the bonus-malus coefficient. As the coefficient increases, due to an accident where the driver is at fault, the marginal cost of a future accident increases. Hence, effort is greater when the premium is higher, or alternatively when the bonus-malus coefficient is higher.

The optimal effort level may also depend on the last contract choice $d_{t-1}$. This first occurs because the presence of switching costs implies that drivers who already have CC coverage are more likely to prefer staying with this coverage than to reduce their coverage and exert more effort on the margin. This type of state-dependence implies that the previous contract choice should be predictive of current accidents. A complementary mechanism to the bonus-malus is at work. Those who had an at-fault accident at $t-1$ incurred different losses depending on whether they were covered or not. This affects disposable income and the marginal utility of consumption. We thus observe, the well-known trade-off between coverage and effort. Because an uncovered loss will lead to a larger marginal utility of consumption than a covered one, the uncovered driver will exert more effort at $t$. These two effects together create a situation where current accident probabilities depend on past contract choice. This form of moral hazard is in the spirit of Lambert
Finally, optimal effort may depend on past accidents through the updating of the subjective belief from $\pi_{t-1}$ to $\pi_t$ following the occurrence of $n_{t-1}$. If the occurrence of $n_{t-1}$ leads to an upward shift in $\pi_t$, a driver may exert more effort to lower his probability of having an accident at $t$. The driver is essentially learning about his level of risk and modifies his behavior accordingly. Thus, learning about risk induces a negative relationship between current accidents and past ones.

The optimal contract choice is a discrete function of the same variables. With asymmetric learning, those who have more accidents and learn that they are more likely to be high risks may choose to purchase more coverage for a given premium. Note however that if the driver is at fault, the premium will increase due to the bonus-malus. Thus, the effect is ambiguous in the case of at-fault accidents. For a given bonus-malus, the prediction is clear for other accidents.

### 3.2. Empirical Predictions

The SOFRES panel contains the following data on individuals $i=1,...,N$ and $t=1,2,3$:

\[
\{\{n_{it}, b_{it}, d_{it}, x_{it}\}_{t=1,2,3}\}_{i=1,...,N}
\]

where $n_{it}$ is equal to 1 if any accident occurs, and zero otherwise (it includes both at-fault and other accidents, either declared or not). The SOFRES panel does not contain information that would allow us to identify at-fault claims; it only tells us whether or not there was a claim for an accident. Both $b_{it}, d_{it}$ have the same definition as in the model and $x_{it}$ is a vector of policyholder-vehicle characteristics (experience, age of vehicle and income).\(^8\) The bonus-malus coefficient at $t$ does not include at-fault accidents since the last renewal of the contract. However, we do not have information on contract renewal dates. Thus we use calendar year to define periods while the model defines those as contract years. This could lead to some measurement error and simultaneity bias, which is mitigated due to the use of lagged variables.

Substituting the solution for optimal effort in the accident process gives us the following expression for accident probability $\Pr(n_{it} | n_{it-1}, d_{it-1}, b_t, \alpha_{it-1}, c_{it}, x_{it})$. We first assume all conditioning variables are observable. Two tests of moral hazard are possible based on the discussion in Section 3.1. The first test involves the bonus-malus coefficient $b_t$. Consequently, our first test is the following:

**Test MH1:** There is no evidence of moral hazard if $b_t$ has no effect on the accident distribution. A negative effect is consistent with the presence of moral hazard.

Because $\pi_{t-1}$ is not observed by the insurer, we cannot confirm the presence of moral hazard by finding a negative effect with the MH1 Test. Theory

---

\(^8\) Many more characteristics of the vehicle and driver are available in the SOFRES survey. However, in preliminary analysis, these were found not to affect the results. Therefore, we opted for a smaller set of characteristics.
would predict that $\pi_{\mu-1}$ would have a negative effect on $n_i$ conditional on $\alpha_i$. Furthermore, since $\pi_{\mu-1}$ is positively correlated with $b_{i\mu}$, this would imply that the negative effect of $b_{i\mu}$ on $n_{i\mu}$ could be explained by learning about risk rather than moral hazard. Finding evidence of a negative effect on MH1 is therefore consistent with moral hazard but does not establish its presence. The fact that $\alpha_i$ is unobserved is problematic because by construction, it is positively correlated with the bonus-malus and with accidents. However, we can account for $\alpha$ as a random effect component provided we deal with the left-censoring problem observed in the data. The problem arises in other tests as well and we will discuss it further in Section 3.3.

The second test involves looking at the relationship between $d_{\mu-1}$ and $n_{\mu}$. Because of switching costs and the fact that optimal effort is higher when the driver was not covered and had an accident, moral hazard would predict a positive effect of past CC coverage on the occurrence of accidents. This leads to a second test for moral hazard:

**Test MH2:** There is no evidence of moral hazard if $d_{\mu-1}$ has no effect on accident probabilities. A positive effect is interpreted as evidence of moral hazard.

The unobservability of $\pi_{\mu-1}$ biases the effect of $d_{\mu-1}$ on $n_{\mu}$ towards zero. This is because $\text{cov}(d_{\mu-1}, \pi_{\mu-1}) > 0$ but $\pi_{\mu-1}$ has a negative effect on $n_{\mu}$ conditional on $\alpha_i$. Accordingly, a positive effect under MH2 implies that moral hazard is present. A zero (or negative effect) does not allow us to confirm the presence of moral hazard due to the presence of asymmetric learning.

As mentioned in the previous section, we can test for asymmetric learning by using the optimal contract choice equation. We can estimate the following contract choice probability from the data $\Pr(d_{\mu} | n_{\mu-1}, d_{\mu-1}, b_{i\mu}, \pi_{\mu-1}, e_{\mu-1}, x_{\mu})$. Given that we control for $b_{i\mu}$, the remaining effect of $n_{\mu-1}$ on $d_{\mu}$ should be positive under asymmetric learning. At-fault claims may also have an effect on contract choice if the premium increases. This will generally not be the case when the driver is not at fault. The fact that $\pi_{\mu-1}$ cannot be observed will only bias the coefficient on $n_{\mu-1}$ upward.

Thus, we propose the following test:

**Test AL:** There is no evidence of asymmetric learning if $n_{\mu-1}$ has no effect on $d_{\mu}$. A positive effect can be interpreted as evidence of asymmetric learning.

The three tests together provide a strong framework for separating moral hazard from asymmetric learning. Only one ambiguous case emerges. It involves finding a negative effect on MH1, no effect on MH2 and a positive effect on AL. From this, we can conclude that there is AL but not that there is MH. This is because rejection of MH1 is consistent with AL due to the fact that $\pi_{\mu-1}$ is unobserved by the insurer. Because MH2 is biased towards zero due to AL, we cannot detect the presence of moral hazard. Table A.4 in the online appendix summarizes the conclusions that can be drawn from these three tests.
3.3. Econometric Model

We build an econometric model that allows us to conduct all three tests jointly. We consider a joint model for the probabilities of accidents and contract choice. We rely on parametric models. We specify each equation as a dynamic binary choice model with predetermined regressors and an error component structure. We let error terms be correlated between equations, as common unobservability of $\pi_{t-1}$ and $\alpha$ would suggest. The error component structure is appealing given the likelihood of serial correlation in contract and accident outcomes. For example, the inability to observe $\alpha$ leads to spurious state-dependence. However, a key question in this context is how to deal with the initial condition problem since histories are left-censored. Below we use the solution proposed by Wooldridge (2005). We first test the adequacy of the model on simulated data.

More specifically, we specify the process for accidents as

$$n_{it} = I(x_{it}\beta_i + \phi_{it}d_{it-1} + \phi_{ity}n_{ity} + \phi_{itb}b_t + \epsilon_{n, it} > 0)$$

where $\epsilon_{n, it}$ has an error component structure $\epsilon_{n, it} = \alpha_{ni} + \nu_{n, it}$. We specify a similar equation for contract choice

$$d_{it} = I(x_{id}i\beta + \phi_{id}d_{it-1} + \phi_{idy}n_{ity} + \phi_{idb}b_t + \epsilon_{d, it} > 0)$$

where again $\epsilon_{d, it} = \alpha_{di} + \nu_{d, it}$. In both (10) and (11), all right-hand side variables are assumed to be predetermined such that they are independent of $\nu_{j, it}, j = d, n$. The bonus-malus variable $b_t$ is a deterministic function of past accidents $b(n_{i0}, ..., n_{it})$. We allow the pair of unobserved heterogeneity terms $\alpha = (\alpha_{ni}, \alpha_{di})$ to be jointly normally distributed with correlation $\rho_{\alpha}$ and variances $\sigma^2_{\alpha, j}, j = n, d$. We do the same for the residual error terms $\nu_{it} = (\nu_{n, it}, \nu_{d, it})$. Conditional on the unobserved heterogeneity terms, omitting the conditioning on $x_{it}$, and writing $(d_{it}, n_{it}) = z_{it}$, we have

$$\Pr(z_{i1}, ..., z_{iT} | \alpha_t, z_{i0}, b_{i1}) = \prod_{t=1}^{T} \Pr(z_{it} | z_{i,t-1}, b_t, \alpha_t)$$

(12)

which is the probability of the joint sequence of accident and contract choices given the initial bonus-malus, contract choice and accident outcome. The period $t = 1$ will generally not be the start of the process. As we already mentioned, we face a problem of left censoring. Equation (12) makes the problem clear. Integrating (12) over $\alpha_t$ implies that we either need to assume that $\alpha_t$ is orthogonal to $z_{i0}, b_{i1}$, that $z_{i0}, b_{i1}$ have a degenerate distribution, or that we know the joint distribution of $(\alpha_t, z_{i0}, b_{i1})$. It is implausible to assume orthogonality and to assume a degenerate distribution. Even for beginners, the bonus-malus appears to vary substantially. Many beginners start with a bonus of 0.5. As a result, we need to specify features of
the joint distribution of \((\alpha_i, z_{i0}, b_{i1})\). Following Wooldridge (2005) we specify \(\Pr(\alpha_i | z_{i0}, b_{i1})\) and maximize likelihood using probabilities in (12) conditional on \(z_{i0}, b_{i1}\). In particular, we assume that \(E(\alpha_{ji} | z_{i0}, b_{i1}) = \pi_{j} z_{i0} + \delta_{j} b_{i1}, j = d, n\). Defining \(\eta_{ji} = \alpha_{ji} - E(\alpha_{ji} | z_{i0}, b_{i1})\) as the residual term, one can substitute \(z_{i0} \pi_{j} + b_{i1} \delta_{j}\) in the index of (10) and (11) and then integrate (12) over the distribution of \(\eta_i = (\eta_{id}, \eta_{in})\). We assume the distribution is bivariate normal. The maximum likelihood estimator for all parameters of the model becomes

\[
\theta_{ML} = \arg \max_{\theta} \frac{1}{N} \sum_{i=1}^{N} \log \int \Pr(z_{i1}, \ldots, z_{iT} | \eta_i, b_{i1}, x_{i1}, \ldots, x_{iT}) d\eta
\]

We compute the two-dimensional integral in (13) by simulation, replacing the integral with the simulator \(R^{-1} \sum_{r=1}^{R} \Pr(z_{i1}, \ldots, z_{iT} | \tilde{\eta}_{ir}, b_{i1}, z_{i0}, x_{i1}, \ldots, x_{iT})\) where \(\{\tilde{\eta}_{ir}\}_{r=1}^{R}\) are draws from the bivariate normal distribution with parameters \((\sigma_{\eta x}^2, \sigma_{\eta d}^2, \rho_{\eta})\). The resulting estimator is the maximum simulated likelihood estimator.\(^9\)

The MH1 test translates into a test of whether \(\phi_{mb} < 0\). A higher bonus-malus, conditional on dynamic selection due to unobserved heterogeneity, gives an incentive to exert more effort and thus reduce accident probabilities. The MH2 test translates into a test of whether coverage in the last period increases accident probabilities in this period; it is a test of whether \(\phi_{md} > 0\) or not. Finally, a test of asymmetric learning is a test of whether an accident last period, conditional on the bonus-malus, leads to an increase in coverage this period. It is a test of whether \(\phi_{dn} > 0\) or not.

3.4. Tests on Simulated Data

The fact that the solution for effort and contract choice is assumed to be linear in the index of (10) and (11) may lead to misspecification. We verify the performance of each test (MH1, MH2 and AL) on simulated data using the model from Section 2. This is particularly useful because some variables, like subjective beliefs and risk type, are unobserved in the data but observed in the simulated data set. We simulate a cohort of drivers and then keep observations from experience years 6 to 15 (10 years). We do not take observations from the start of the process precisely so that we have left-censoring, i.e. initial outcomes at year 5 are correlated with risk type. This replicates what we have seen in the data. We stop at year 15 because few contract changes occur after this date and we already have 10 observations per driver used in the estimation. We include income and experience (linear) as controls in the vector \(x_{it}\).\(^10\)

---

\(^9\) We use 50 draws along each dimension. We use the BFGS numerical maximization algorithm and compute robust standard errors with the sandwich estimator.

\(^10\) We use 50 draws along each dimension. We use the BFGS numerical maximization algorithm and compute robust standard errors with the sandwich estimator.
We consider four scenarios where we modify the information problems present in simulated data. The top panel of Table 2 report the results of the three tests for all four scenarios.

**TABLE 2. Empirical tests on simulated data.**

<table>
<thead>
<tr>
<th>Underlying DGP</th>
<th>MH</th>
<th>No MH</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Test on Contract Choice</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(n(t-1) - AL)</td>
<td>0.428</td>
<td>-0.062</td>
</tr>
<tr>
<td>(0.001)</td>
<td>(0.327)</td>
<td>(0.001)</td>
</tr>
<tr>
<td><strong>Test on Accidents</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(b(t) - MH1)</td>
<td>-0.575</td>
<td>-0.674</td>
</tr>
<tr>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.537)</td>
</tr>
<tr>
<td>(d(t-1) - MH2)</td>
<td>0.116</td>
<td>0.1344</td>
</tr>
<tr>
<td>(0.003)</td>
<td>(0.001)</td>
<td>(0.013)</td>
</tr>
</tbody>
</table>

| Specification | | Heterogeneity Risk |
|---------------|--------------------------|
| **Test on Contract Choice** | | |
| \(n(t-1) - AL\) | 0.456 | 0.421 |
| (0.001) | (0.001) | (0.001) |
| **Test on Accidents** | | |
| \(b(t) - MH1\) | 0.109 | -0.424 |
| (0.159) | (0.001) | (0.001) |
| \(d(t-1) - MH2\) | 0.176 | 0.095 |
| (0.001) | (0.001) | (0.001) |

Notes: The table reports bivariate probit coefficient estimates along with p-values. The top panel presents estimates obtained by maximum simulated likelihood with 50 draws per respondent and equation using the solution proposed by Wooldridge (2005) for the initial condition problem. Different specifications are based on simulated data from varying data generating processes (DGP). The first column assumes both moral hazard (MH) and asymmetric learning (AL). The other three specifications vary the presence of information problems. Each specification controls for experience and income and deals with the initial condition problem as mentioned in the text. The bottom panel performs various robustness checks. The first specification does not control for unobserved heterogeneity nor corrects for left censoring. The second specification controls directly for the risk type. The third specification adds the lag subjective probability of being high risk type. The last column adds heterogeneity in risk aversion, negatively correlated with unobserved heterogeneity in risk type. Complete results available upon request.

First, we assume both moral hazard and asymmetric learning are present. All three tests capture the information problems. This is shown in column 1 of the top panel. Lag accidents have a positive and statistically significant effect on the procurement of CC coverage as predicted by asymmetric learning conditional on the bonus-malus. Both moral hazard tests, MH1 and MH2, reveal the presence of moral hazard. The current bonus-malus has a negative effect on accidents while the lag of the contract choice has a positive effect. Column 2 reports what happens when we assume symmetric learning rather than asymmetric learning. Both insurers and drivers observe accidents. In that case, the AL test does not pick up a residual effect.
of lag accidents on contract choice. This implies that within the context of the model the bonus-malus is as good as the subjective belief of the driver and corrects the premium such that no increase in coverage occurs. The third column shows what happens when we simulate data with asymmetric learning but no moral hazard. In that case the AL test picks up asymmetric learning while both MH tests do not identify moral hazard. The MH2 test even shows that lag contract choice is negatively correlated with accidents. The last column shows the results when we assume that there is no information problem; in that case, all tests yield negative conclusions. The AL test yields a negative effect of lagged accidents on current contract choice. Overall, the tests on simulated data appear to pick up the information problems assumed in the model despite the restrictive functional form assumed and the unobservability of some of the state variables.

Next, we verified whether our solution to control for unobserved heterogeneity and initial conditions is adequate. Although results in Table 3 are generally positive, we demonstrate what happens if we control directly for the risk type (because in the simulated data we know who is high risk), if we do not control for the risk type, if we add the lag subjective probability \( \pi_{t-1} \) as a control and finally if we allow for heterogeneity in risk aversion. We report the results of the tests in the bottom panel of Table 2 using data generated with moral hazard and asymmetric learning present in the DGP.

The first column shows what happens if we do not account for initial conditions and unobserved heterogeneity. For the AL and MH2 tests, results are relatively robust. However, the MH1 test reveals an insignificant effect of the bonus-malus on accidents despite the presence of moral hazard in the data. This is because the bonus-malus is significantly positively related to risk type. If the bonus-malus is omitted, we have an upward bias in the coefficient as suggested by Abbring et al. (2003).

The second column reports results where we control for the risk type \( \alpha_i \) directly. In this case, we do not need to control for the initial conditions \( z_{i0}, b_1 \) either. The results are quite close to those in column 1 of the top panel of Table 2, suggesting that our method for controlling for initial condition and unobserved heterogeneity is adequate. Because \( \pi_{t-1} \) is also unobserved, we produce results in column 3 where we include it as a control. Again the conclusions do not change.

The last check we do is to allow for heterogeneity in risk aversion. We allow for half the respondents to have a risk aversion parameter 1.35 while the other half has 1.65. Risk aversion is known to drivers. To consider the case raised by Finkelstein and McGarry (2006) where the more risk averse are also the less risky, we allow \( \delta_H \) to be higher for those with low risk aversion. We set \( \delta_H = 0.2 \) for the low risk aversion types and \( \delta_H = 0.4 \) for the high risk aversion types. We simulate

11. We analyzed the sensitivity of our results to underlying structural parameters of the DGP. In Table A.5 of the online appendix, we show how varying risk aversion and the productivity of effort in reducing accidents changes the result. Lowering the productivity of effort tends to reduce our ability to detect moral hazard under MH1 and MH2. The asymmetric learning test is weaker when risk aversion is larger, although remaining statistically significant. This is in part due to the fact that there are fewer transitions in contract choice over time.
1000 drivers of each type (so that sample size is the same as in top panel of Table 2). The results are presented in the last column of the bottom panel of Table 2. The results of the AL and MH1 tests are robust to heterogeneity in risk aversion, but the ability to detect moral hazard through MH2 is weakened when we introduce heterogeneity in risk aversion. This is because heterogeneity in risk aversion is not an additive fixed effect in the theoretical model. Rather, the response of effort to contract parameters is affected in a non-linear way by risk aversion.

4. Results on SOFRES Panel

We now apply our tests to the SOFRES panel. We estimate separate models for two groups: those with less than 15 years and more than 15 years of experience in the first survey year considered. We consider the first sample (less than 15 years) the “inexperienced” sample and the second (more than 15 years) the experienced sample. One might suspect that asymmetric learning is more likely among drivers with less experience. This may be the result of the bonus-malus scheme because in the long-term the coefficient will capture unobserved risk types through at-fault accidents. Table A.3 in the online appendix gives descriptive statistics on transitions in the panel stratified by experience.

In Table 3, we report a summary of the estimation results for the two groups (complete results are available upon request). We also consider a specification where we allow for a separate MH2 and AL effect for drivers with less than 5 years of experience (the beginners). We do not find evidence of learning about risk among the inexperienced group overall. The point estimate on lag accidents in the contract equation is positive but statistically insignificant (column 1). However, when we allow for a different asymmetric learning effect for the beginners, we find a large positive effect (0.863) that is barely statistically significant at the 5% level ($p$-value = 0.053, column 3). We do not find evidence of asymmetric learning among more experienced drivers (column 2). The point estimate is negative (-0.223) and statistically insignificant ($p$-value=0.483). This suggests that asymmetric learning vanishes relatively quickly.

The MH1 and MH2 tests generally agree for both groups. We find evidence of moral hazard among inexperienced drivers (less than 15 years, column 1). The point estimate is strongly negative on the bonus-malus (-2.241, $p$-value=0.081). This suggests that accidents at fault at $t-1$ trigger incentives to exert caution, resulting in fewer accidents at $t$. The MH2 test tells a similar story. The point estimate is positive (0.651) and statistically significant ($p$-value=0.026), as the theory would predict. We do not find stronger evidence of moral hazard among beginners (column 2). The effect for those with less than 5 years of experience is no different from that of the rest of the group (point estimate is 0.17, $p$-value=0.351). There is no evidence of moral hazard among the more experienced group. The point estimate is 0.096 on the MH1 test with a $p$-value of 0.892. The same conclusion is reached using the MH2 test.
5. Conclusion

In this paper, we analyze the identification of moral hazard from asymmetric learning about risk within the context of a structural dynamic insurance model. We extend the model developed by Abbring et al. (2003) to include learning about risk and insurance coverage choice. We derive two tests in addition to their negative occurrence dependence test, which we apply on longitudinal data from France for the period 1995-1997.

Despite the short horizon of the panel, our results suggest the presence of moral hazard among inexperienced drivers with less than 15 years of experience. We do not find evidence of asymmetric learning for the vast majority of drivers. We find some evidence of asymmetric learning for those with less than 5 years of experience, which disappears quickly as both drivers and insurers learn about the underlying risks. The results for the experienced group are largely consistent with the evidence presented in Abbring et al. (2003).

### TABLE 3. Tests on SOFRES panel.

<table>
<thead>
<tr>
<th>Test on Contract Choice</th>
<th>Years of Experience Group</th>
<th>Interaction &lt; 5</th>
<th>More than 15</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Less than 15 yrs</td>
<td>yrs</td>
<td>years</td>
</tr>
<tr>
<td>n(t-1) - AL Test</td>
<td>0.152</td>
<td>0.034</td>
<td>-0.223</td>
</tr>
<tr>
<td></td>
<td>(0.729)</td>
<td>(0.940)</td>
<td>(0.483)</td>
</tr>
<tr>
<td>n(t-1) x exp&lt;5</td>
<td>0.868</td>
<td>(0.053)</td>
<td></td>
</tr>
<tr>
<td>Test on Accidents</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b(t-1) - MH1 Test</td>
<td>-2.241</td>
<td>-2.216</td>
<td>0.096</td>
</tr>
<tr>
<td></td>
<td>(0.081)</td>
<td>(0.080)</td>
<td>(0.892)</td>
</tr>
<tr>
<td>d(t-1) - MH2 Test</td>
<td>0.651</td>
<td>0.622</td>
<td>-0.14</td>
</tr>
<tr>
<td></td>
<td>(0.026)</td>
<td>(0.033)</td>
<td>(0.661)</td>
</tr>
<tr>
<td>d(t-1) x exp&lt;5</td>
<td>0.170</td>
<td>(0.351)</td>
<td></td>
</tr>
</tbody>
</table>

Sample Size (N x T) 1066 1066 1537

Notes: This table reports bivariate probit coefficient estimates along with p-values based on robust standard errors. Estimates obtained from maximum simulated likelihood with 100 draws per respondent and equation. The first column reports coefficients estimated on drivers with less than 15 years of experience. The second column allows for different AL and MH2 tests for those with less than 5 years of experience in the contract choice equation. The last column presents the coefficients estimated only on those with more than 15 years of experience. Other controls include experience, income and age of vehicle as well as initial conditions. Complete results available upon request.
References


Journal of Economics, 14, 441-452.