Does Experience Rating Matter in Reducing Accident Probabilities? A Test for Moral Hazard

Olivia Ceccarini
University of Pennsylvania

November, 2007

Abstract

I examine the empirical importance of moral hazard, adverse selection and state dependence in the determination of motor vehicle accidents, using a unique longitudinal dataset on car insurance policies that I obtained from an Italian insurance company. I develop a dynamic model of driving effort given an insurance contract that allows for moral hazard, for adverse selection, for state dependence in accidents and for general premium pricing schedules. Simulations based on a calibrated version of the structural model suggest that driving effort increases in the marginal premium price increase, which is observed in the data. From the model, I derive an approximate decision rule for driving effort that underlies my empirical specification for the probability of having an accident. The estimating equation is a dynamic discrete choice panel data model with unobserved heterogeneity, state dependence and predetermined variables. My unusually rich data permit a flexible way of incorporating all of these factors in the model. I estimate the model using the semi-parametric estimator of Arellano and Carrasco (2003). I find much stronger evidence of moral hazard than in the previous literature. In particular, there is negative dependence between accident probabilities and the experience rating class. Thus, monetary costs are important determinants of accident probabilities. There is also evidence of negative state dependence suggesting that drivers who recently had accidents engage in accident avoidance behavior.

1 Introduction

Informational asymmetries are potentially important to the operation of insurance markets. They take the form of information that is known to the insured but not to the insurer that is relevant to the risk being insured. Interest in identifying the presence of information asymmetries goes beyond curiosity, as they can lead to market inefficiencies and possible market breakdowns. In some cases, they justify government interventions, such as making insurance mandatory (Akerlof, 1970). Information asymmetries can be divided into two main types: adverse selection and moral hazard.

Static models of insurance contracts predict that in the presence of information asymmetries there will be a positive correlation between the probability of having an accident and the level of insurance bought, which can be measured by the choice of deductible. In the presence of adverse selection, the positive correlation is attributable to some characteristic that the policy holder knows about himself, such as driving ability, which is unknown to the insurance company. Individuals with worse than average driving abilities typically self-select into contracts with lower deductibles. Moral hazard occurs if individuals change their behavior in response to the features of the insurance contract. Contracts with higher insurance coverage and, hence a low deductible, make the cost of an accident less expensive. Under moral hazard, the insured drives less cautiously when the level of insurance is higher.

Interestingly, the positive correlation between the probability of having an accident and the level of insurance, which is so widely accepted in theory, has proved hard to establish empirically. For example, Chiappori and Salanie (2000) test for the presence of asymmetric information using a cross section of French contracts. The authors employ a bivariate probit where the outcomes are the choice of deductible and whether an accident (a claim) occurs. They show that if asymmetric information is present, the two error terms should be positively correlated. The authors find no evidence of asymmetric information. A more recent paper by Cohen (2005) tests for the presence of asymmetric information on a sample of Israeli automobile policyholders employing the same approach as used by Chiappori and Salanie. She finds some evidence for asymmetric information, particularly for more experienced drivers.

A drawback of the test performed in both studies is that it does not disentangle moral hazard from adverse selection and the test could break down under certain kinds of unobservable heterogeneity. For example, individuals at low risk for accidents could also be more risk averse and buy a larger amount of insurance. Failure to control for this unobserved heterogeneity among drivers (risk preferences) could make it difficult to find evidence of moral hazard even when present in the data.

These considerations suggest the value of looking at repeated observations on the relationship between policyholders and insurance companies. With longitudinal data, a single policy holder is observed at different times facing different marginal price increases of an accident. Within policyholder variation in behavior over time allows one to identify moral hazard taking into account unobserved heterogeneity.

The recent theoretical literature has studied the properties of long term dynamic insurance contracts and one important finding is that the pricing schemes depend on the source of information asymmetry. With adverse selection, optimal contracts are characterized by pricing schemes that depend less on the claim history (Hayes and Cooper, 1987; Rubinstein and Yaari 1987). Another general finding from that literature is that long term contracts display very complicated schemes, schemes that are not observed in actual auto insurance markets. In most countries, auto insurance contracts exhibit fairly simple experience rating formulas, in which the history of claims affects tomorrow's premium, that is, if an insured has an accident today, tomorrow's premium will increase and vice versa.

Because observed insurance contracts are simpler than those predicted by dynamic insurance models, most of the recent empirical literature takes the feature of the contracts as given and analyzes how drivers react to incentives. Abbring, Chiappori and Pinquet (2003) develop a continuous time model for accidents in which, at any given time, the driver chooses a probability of having an accident and is subject to an experience rating premium pricing scheme. The authors show that under moral hazard, the experience rating affects the agent's behavior. That is, the occurrence of a claim affects the entire discounted cost of future claims making them more expensive, and the agent will react by exerting more cautious behavior in the following period. The model implies a simple testable prediction: negative state dependence in claims. The authors test for negative state dependence using a duration model estimated on data on French insurance contracts. They find some evidence of moral hazard although the estimate is not statistically significant. Another more recent study by Abbring, Chiappori and Zavadil (2007) extends the model to allow for more general premium schedules and proposes a different test for moral hazard that examines whether the duration of time between insurance claims depends on the experience rating class. Using data from a Dutch insurance company, the authors find some evidence of moral hazard. A study by Israel (2004) also tests for dependence of the accident probability on the experience rating class using a random effect estimator applied to longitudinal data from the U.S. and finds a small but statistically significant dependence. All of the tests employed in these studies do not allow for true state dependence. For example, an insured that experiences a claim could be just less prone to have an accident in the following period only because of the recent memory of having an accident, irrespective of any premium changes. Also, the random effect estimator used in the Israel (2004) study imposes strong and arguably implausible assumptions on the form of adverse selection.

This paper investigates the empirical importance of moral hazard, adverse selection and state dependence in the determination of motor vehicle accidents, using a new longitudinal data set on car insurance policies. I obtained the data from an Italian insurance company's administrative records. The data set follows policyholders over the period 2000-2005. The first part of the paper develops a theoretical model of driving effort that combines features of the models of Abbring, Chiappori and Pinquet (2003) and Abbring, Chiappori and Zavadil (2007), but also incorporates state dependence and adverse selection. My strategy for disentangling moral hazard from adverse selection makes use of the fact that the same policyholder faces different price increases of an accident at different times. I can also isolate the effect of true state dependence, because drivers in the same experience rating class have different recent accident histories. The theoretical model has the prediction that in the presence of moral hazard, drivers would take into account the marginal price increase of an accident in choosing their driving effort. Simulations from a calibrated structural model suggest that, under certain parameterizations, driving effort is increasing in the marginal price increase of an accident.

From the model, I derive an approximate decision rule for driving effort that underlies my empirical specification for the probability of having an accident. The estimating equation is a discrete choice dynamic panel data model with unobserved heterogeneity, state dependence and predetermined variables. I estimate the model using a semi-parametric estimator proposed in Arellano and Carrasco (2003), which allows for correlated random effects with a non-parametrically specified mean as well as predetermined variables. My empirical results show much stronger evidence for moral hazard than has been demonstrated in the earlier literature. In particular, I find negative dependence between accident probabilities and the experience rating class, with higher marginal increases in insurance *premia* at higher classes associated with larger decreases in probabilities. Thus, monetary costs are important determinants of accident probabilities. There is also evidence of negative state dependence, after controlling for adverse selection and moral hazard, suggesting that drivers who recently had accidents engage in accident avoidance behavior.

This paper is organized as follows. Section 2 presents the model and its testable implications. Section 3 introduces a numerical example. Section 4 describes the data set and the experience rating system. Section 5 discusses the estimation approach, which makes use of a recently developed estimator by Arellano and Carrasco (2003). Section 6 reports the results from the main estimator and shows the differences when compared with the ones derived by adopting random effect probit estimators and fixed effect logit estimators. Section 7 addresses potential sample selection concerns. Section 8 concludes.

2 The Driver's problem

Time is measured in contractual years and the process begins when the driver enters in the insurance company for the first time. The driver chooses an effort level $e_t \in [0, \bar{e}]$. Let us define the probability that an accident does not occur between t and $t+1$ as $p(e_t, \eta, a_{t-1})$. The probability depends on effort, on risk level η, which is related to the ability to drive, and on the random variable a_{t-1}, that takes value 1 if the insured had an accident between time $t-1$ and t and 0 otherwise. We assume that p is increasing, concave and differentiable in e. These assumptions mean that the probability of having an accident is decreasing in effort and that there is diminishing marginal return to effort. The $p(e, a, \eta)$ function is is assumed to be decreasing in η, conditional on e_t and a_{t-1}, which means that individuals with low risk are less prone to have an accident. We interpret $p(e, \eta, 1) - p(e, \eta, 0)$ as the direct effect of having had a recent accident on the current accident probability. For example, a recent accident might make one more cautious or more aware, in which case $p(e, \eta, 1) - p(e, \eta, 0) > 0$.

At any given time t, the individual is endowed with some wealth w_t. The premium that the insured pays is denoted by π. The premium schedule is given by a convex function $\pi : K \longmapsto R_+$ where $k = 1....K$ is the set of the experience rating categories. The experience rating category at time $t+1$, which we define as k_{t+1}, depends on the previous experience category and whether or not the insured had an accident, $k_{t+1} = f(k_t, a_t)$. The utility cost of exerting effort is given by $c(e)$. The cost function c is twice differentiable, increasing and convex. If the contract does not include a deductible the agent's instantaneous payoff is given by $u(w_t - \pi_t(k_t)) - c(e_t))$ where u is twice differentiable, increasing and concave.

The agent maximization problem is the one of choosing e_t to maximize the total expected utility:

$$\sum_{t=1}^{\infty} E[\delta^t(u(w_t - \pi(k_t)) - c(e))].$$

We can rewrite the problem in the recursive form:

$$\begin{aligned} V(k, \eta, a) \quad = \quad & \max_e u(w - \pi(k)) - c(e) + \delta\{p(e, \eta, a)V(f(k, 0), \eta,\ 0) + \\ & (1 - p(e, \eta, a))V(f(k, 1), \eta, 1)\}. \end{aligned}$$

Because the objective function on the right hand side is concave in e, optimality, for interior solutions, requires:

$$c'(e^*) = \delta\{\frac{\partial p(e^*, \eta, a)}{\partial e} V(f(k,0), \eta, 0) - \frac{\partial p(e^*, \eta, a)}{\partial e} V(f(k,1), \eta, 1)\},$$

which can be rewritten in the following form:

$$\frac{c'(e^*)}{\frac{\partial p(e^*, \eta, a)}{\partial e}} = \delta\{V(f(k,0), \eta, 0) - V(f(k,1), \eta, 1)\},$$

from which we obtain an optimal decision rule for effort:

$$e^*(k, \eta, a).$$

In absence of moral hazard the effort level e does not depend on k, so that the probability of not having an accident depends only on the risk level of the individual η and on whether or not the policy holder had an accident. To show that in the presence of moral hazard the effort level is increasing in k, it would be enough to show that the difference $V(f(k,0), \eta, 0) - V(f(k,1), \eta, 1)$

is increasing in k, which would follow if the value function were concave in k. By standard dynamic programming argument it is enough to show that for every continuous $g(k, \eta, a)$ that is concave in k, the objective function $u(w - \pi(k)) - c(e) + \beta\{p(e, \eta, a)g(f(k, 0), \eta, 0)) + (1 - p(e, \eta, a))g(f(k, 1), \eta, 1\}$ is jointly concave in (e, k, a). However, this sufficient condition can fail to hold so that the standard approach does not apply. Nevertheless, we can explore which assumptions on the functional forms for the utility function, cost function, and premium would imply that the effort level rises in the marginal cost of an accident, which is a finding from the data analysis.

To summarize, the above function yields an optimal choice of effort as a function of accident history, experience class, and unobserved heterogeneity. Plugging in the optimal effort choice gives the optimal probability:

$$p(e^*(k, \eta, a), \eta, a).$$

In the empirical work reported below, we estimate a reduced form approximation to this probability function.

3 An Example

In this section we solve a discrete choice simplified version of the model presented in the previous section. It abstracts from state dependence and unobserved heterogeneity and focuses only on the choice of effort as it relates to accident history. The difficulties in showing that V is concave do not depend on the horizon, on the unobserved heterogeneity or on the presence of state dependence so here we present a simpler model. Because there is a mapping between the effort choice and the probability, I assume that the agent chooses a probability p of not having an accident. I use the model to explore whether, for reasonable values of the parameters, the probability of not having an accident is increasing in the experience rating class, as I observe in the actual data. I define k_{t-1} as the experience class coming in to period t, p_t is the probability of not having an accident at time t, chosen by the decision maker through the choice of driving effort and $a_t = 1$ if the policy holder has an accident in period t, else $= 0$.

It is assumed that drivers make decisions about driving effort, which in turn determines the probability of not having an accident, in period $1 ... T$. The dynamic model can be described in Bellman equation form:

$$V_t(k_{t-1}) = \max_{p_t} U(w - \pi(k_{t-1})) - c(p_t) + \delta E(V_{t+1}(k_t)).$$

In the last decision period, the value function is:

$$V_T(k_{T-1}) = \max_{p_T} U(w - \pi(k_{T-1})) - c(p_T).$$

The law of motion for the experience class is

$$k_t = k_{t-1} + a_t,$$

so that experience class increases by 1 if the driver has an accident and otherwise stays the same.

The derivations will use the following assumed functional forms for the utility function and cost function. The utility function is assumed to be linear and the cost function convex.

$$\begin{aligned} U(w - \pi(k_{t-1})) &= w - \pi(k_{t-1}) \\ c(p_t) &= p_t^{\beta}, \quad \beta > 2. \end{aligned}$$

Under the assumed functional form, the cost is zero if $p_t = 0$ (i.e. no effort). The parameters are set at the following values: $\beta = 5, w = 50000, \delta = 0.97$. The premium is assumed to be equal to $(\frac{k}{10})^3$, where k is an experience class which is the total number of previous accidents.

3.1 Model solution

First, consider the solution in the last period T. Because effort is costly and the only gain is reduced premium in the future, the solution in the last period will be to apply no effort, so $p_T^* = 0$. The value function in the last period is therefore:

$$w - \pi(k_{T-1}).$$

Next, consider the second-to-last period $(T-1)$. Plugging in the value function for the last period, the individual solves the problem:

$$\begin{aligned} & \max_{p_{T-1}} w - \pi(k_{T-2}) - p_{T-1}^{\beta} + p_{T-1}\delta(w - \pi(k_{T-1})) + \\ & (1 - p_{T-1})\delta(w - \pi(k_{T-1} + 1)) \\ = \; & \max_{p_{T-1}} w - \pi(k_{T-2}) - p_{T-1}^{\beta} + p_{T-1}\delta(w - \pi(k_{T-2})) + \\ & (1 - p_{T-1})\delta(w - \pi(k_{T-2} + 1)). \end{aligned}$$

Taking the first order conditions we obtain:

$$\begin{aligned} & -\beta p_{T-1}^{\beta-1} + \delta(w - \pi(k_{T-2})) - \delta(w - \pi(k_{T-2}+1)) \\ = \; & -\beta p_{T-1}^{\beta-1} + \delta(\pi(k_{T-2}+1) - \pi(k_{T-2})) = 0, \end{aligned}$$

so

$$\beta p_{T-1}^{\beta-1} = \delta(\pi(k_{T-2}+1) - \pi(k_{T-2})),$$

which implies that for an interior solution (apply some driving effort):

$$p_{T-1}^{*} = \left(\frac{\delta[\pi(k_{T-2}+1) - \pi(k_{T-2})]}{\beta}\right)^{1/(\beta-1)}.$$

This expression shows that the driving effort applied to not have an accident will be increasing in the marginal increase in the premium. If the solution is not interior, then $p_{T-1}^{*} = 0$ or $p_{T-1}^{*} = 1$.

The optimal value for p_{T-1}^{*}, the value function in period $T-1$ is given by:

$$\begin{aligned} V_{T-1}(k_{T-2}) \; = \; & \max_{p_{T-1}} w - \pi(k_{T-2}) - ((p_{T-1}^{*})^{\beta}) + p_{T-1}^{*}\delta(w - \pi(k_{T-2})) + \\ & (1 - p_{T-1}^{*})\delta(w - \pi(k_{T-2}+1)). \end{aligned}$$

In period $T-2$,

$$\begin{aligned} V_{T-2}(k_{T-3}) \; = \; & \max_{p_{T-2}} w - \pi(k_{T-3}) - p_{T-2}^{\beta} + p_{T-2}\delta V_{T-1}(k_{T-3}) + \\ & (1 - p_{T-2})\delta V_{T-1}(k_{T-3}+1). \end{aligned}$$

Taking the first order conditions we obtain:

$$-\beta p_{T-2}^{\beta-1} + \delta[V_{T-1}(k_{T-3}) - V_{T-1}(k_{T-3}+1)] = 0,$$

which implies

$$p_{T-2}^{*} = \left(\frac{\delta[V_{T-1}(k_{T-3}) - V_{T-1}(k_{T-3}+1)]}{\beta}\right)^{1/(\beta-1)}.$$

The same expression will hold for periods $T-3$ to 1.

Figure 1 plots the life cycle probability of not having an accident for a representative driver. The optimal probability choice is lower for younger drivers who face a low marginal cost of an accident (they start in the lowest experience class). Each time they have an accident and therefore enter into a higher experience class, the optimal probability choice increases discretely. In time periods where there is no accident, the optimal probability declines slightly due to the horizon effect. Accidents are more costly when there is a longer horizon, since all premium increases are permanent. The horizon effect also counts for the decrease in the optimal probability near the end of (driving) life.

Figure 2 plots the relationship between the optimal probability choice and the experience class for a given time period, in other words, holding constant any horizon effect. The figure is generating by simulating 50 accident histories under the model. The figure shows that the choice of the probability of not having an accident is increasing in experience class, which suggests that drivers exert more driving effort when the marginal cost of an accident is higher.

4 Data Description

4.1 The Experience Rating System

Most insurance companies base their pricing policies on 2 factors: the base premium and the experience rating class. The base premium is usually evaluated considering characteristics of the insured, of the vehicle and of the amount coverage bought. For example, base *premia* differ by geographic location, by gender and by age.

The experience rating class is assigned according to the number of claims that the insured had in the contractual year. In particular, the insurance company provides a discount on the base premium if a claim did not occur in the current period or otherwise adds an extra charge if the policyholder made a claim. Base premia are used to classify the insured into his respective class of risk, and hence they are usually considered as a tool to distinguish between high and low risk individuals. Experience rating schemes, on the other hand, have multi-fold objectives. As described by Lemaire (1995), they introduce an a-posteriori measure of the real risk of the policyholder by linking each single individual to his history of claims. In this respect they play the same role as the base premium, even though they should be related to the class to which the individuals belong to be considered actuarially fair. The second function of the experience rating is to lessen the effect of moral hazard effect. If the marginal price increase in the *premia* is increasing in the experience rating class additional claims become more expensive therefore introducing an incentive to exert effort.

In the Italian insurance market, each company is free to decide the pricing system they want to adopt, according to their own cost structure and past driving records of their portfolio. Both base premia and experience rating classes are freely set by each insurance company, but their implementation is subject to prior approval by ISVAP, the supervisory authority.

Insurance contracts are made without commitment. At each renewal date, either the insurer or the insured can decide to end the contract. However, because in Italy insurance is mandatory, the policyholders can re-apply with the same insurance company and be re-enrolled. A peculiarity of the insurance contracts of this company is that they do not have a deductible. The actuarial office of the insurance company explained that this peculiarity is due to difficulties in recouping the value of the deductible once a claim occurs. In fact, the insurance company under study used to pay claim in advance expenses and ask for a refund from the policy holders. This practice was abandoned since customers would default in paying with a consequent loss.

The discount on the base premium that a policyholder gets depends on the experience rating class. There are 26 experience rating classes with the lower classes having the lower *premia*. A person that enters for the first time in insurance is placed in class 14. Policyholders coming from another insurance company must provide a certificate showing their driving records according to which they will be placed in a given class. If a policyholder does not experience a claim in the current year he moves down a class, if instead he experiences a claim he will move up one or two classes depending on the experience class top which he belongs. For example, a policyholder in class 1 who experiences a claim in the current year will be placed in class 2A moving up only 1 class, whereas a policyholder belonging to class 3 who experiences exactly one claim will be placed in class 5 in the following year. Table 1 shows the experience rating system adopted by the company in more detail, taking into account the fact that the policyholder might experience more than a claim in a given year

The pricing scheme in my data set differs from the proportional pricing scheme studied by Abbring, Chiappori and Pinquet (2003). In Abbring, Chiappori and Pinquet (2003) the premium paid by the policyholder subsequent periods depends on the today's premium in the following way. If the policyholder does not incur in an accident the future premium is equal to 95% of the current premium, whereas if the he incurs in an accident the future premium is 1.25 times the current one. In the case of my data, the experience rating scheme adopted is not proportional. For example, a policyholder in class 8 who experiences a claim in the current period will move up to class 10 with a 11% increase with respect to the current premium, on the other hand a policyholder who is in class 9 and experiences a claim will be moved to class 11 with an 13% increase. The third column of Table 2 shows the percentage increase with respect to the current premium by class. Percentage decreases in the premium are not proportional as well. For example a customer in class 8 that experiences a claim free year will pay move down to class 7 and pay 95 per cent of the current premium whereas a customer in class 12 who does not incur in a claim will move to class 11 and pay 94 per cent of the current premium.

4.2 The Data on Policyholders

The data set consists of longitudinal data coming from the records of a medium size Italian insurance company over a five year period. The data include information on the insurer's operation from May 2000 to May 2005, namely details on the characteristics of the contracts (value of the premium, experience rating class, exposition to risk, risk covered), policy-holder demographics (gender, province of residence, age), description of the vehicle (type of vehicle, horse power) and realization of risk covered (date of the accident, number of claims, claim sizes and whether the policy-holder was at fault or not). The original data consisted of more than 700,000 records regarding all the non-life activities of the company, such as insurance against theft and fire. I limit the sample to auto-insurance records, I ignore cars that were owned by companies or under leasing. The panel is unbalanced in the sense that not all contracts remain in the insurance company's portfolio for the whole period of observation. The resulting data contain information pertaining to about 300,000 policy-holders.

For computational reasons that necessitate using a smaller sample, I focus on a sample of policyholders that enter the company portfolio in year 2000 and remain in the sample for the whole period of observation. Table 3 describes the distribution of contracts by experience rating class in each single period. The majority of the contracts is class 1A in the first period, 1B in the second third and fourth period and 1C in the fifth period[1]. Table 4 gives the distribution of contracts by number of claims. Claims are not rare, around 6 per cent of the sample experiences at least one claim in each single year, however contracts with multiple claims are unlikely.

The second column of Table 2 shows the marginal price increases according to class in case of a claim. The marginal price increase is the difference between the price that the policyholder would pay in case that an accident occurs and the price that he would pay if an accident does not occur. The marginal price increases depend and the experience class with a mostly monotone behavior. The marginal price increase is weakly increasing with the experience rating class up to class 1, jumps down at class 1, and then follows an almost monotone behavior as the experience rating class increases. This means that people belonging to different experience rating classes face different marginal price increases and this difference in incentives should map in different level of efforts with people who face a higher marginal price increase exerting a higher level of effort and, therefore, experiencing a lower probability of having an accident. This simple observation allows us to test for the presence of moral hazard as negative dependence between the experience rating class and the probability of having claim. The theoretical model predicts only dependence between the experience rating class and the accident probability but this insight, which is confirmed by the numerical example, provides an additional testable implication. We require the parameter to be not only different from zero but also negative. In sum, we have two tests for moral hazard, around the null, that is no moral hazard, the parameters associated to the experience rating system should all be zero in both tests. What makes the two tests different is the value of the parameters of the experience rating class. In the first test we require the parameters to be jointly different from zero, in the second test we require the parameters of the experience rating class to display a decreasing pattern.

[1]According to the insurance company files new entrants can not be placed in class 1B but our records indicate that 34 individuals who are new customers are placed in class 1B. I decided not to ignore these customers as this placement might be due to particular sales practices implemented by the agent or to the fact that maybe the customer is a dependent of the company getting an especially good deal.

5 The Econometric Model

The theoretical model of section two implied a decision rule for driving effort, which in turn determines the probability of having an accident. In this section, I estimate a reduced form model for the accident probability as a function of accident history, experience class, and unobserved heterogeneity. A test for moral hazard is a test for whether the accident probability depends on the experience class. The estimating equation is a discrete choice dynamic panel data model.

Dynamic linear panel data models with unobserved effects have been widely studied in the literature. The unobserved effects can be eliminated by first differencing and using appropriate GMM techniques (Anderson and Hsiao, 1982; Arellano and Bond 1991). The literature on non linear dynamic models is less developed but some estimators are available. For example, Honore and Kiriziadou (2000) propose an estimator that allows for state dependence but requires the other regressors to be strictly exogenous[2]. Honore and Lewbel (2002) propose an estimator that allows for state dependence and predetermined variables but requires the existence of an absolutely continuous regressors that is strictly exogenous (a so-called "special regressor").

In this paper, I employ an estimator developed by Arellano and Carrasco (2003), which generalizes the correlated random effects estimator of Chamberlain (1980). Arellano and Carrasco propose a GMM estimator that is more flexible in that it specifies the mean of the distribution of the unobserved heterogeneity non-parametrically.

From the theoretical model we derived the following equation

$$y_{it} = 1\{\alpha + \beta y_{it-1} + \sum_{j=2}^{J} \gamma_j k_{jit} + \eta_i + u_{it} \geqslant 0\},$$

where y_{it-1} is a binary variable taking value one if the insured had an accident between time $t-1$ and t, k_{jit} is a set of predetermined variables that takes value 1 when individual i belongs to the experience rating class j at time t and zero otherwise, η_i represents the driver's risk which is unobserved to the econometrician and u_{it} represents the idiosyncratic error term. $u_{it} + \eta_i | w_i^t \sim N(E[\eta_i | w_i^t], \sigma_t)$[3] where $w_i^t = \{w_{i1} .... w_{it}\}$ with $w_{it} = (\ y_{it-1}, k_{2it} ... k_{Jit})$. By

[2] In a more recent paper Honore and D 'Addio (2006) extend the model to allow for and additonal lag of the dependent variable. However, the generalization to generic regressors that are functions of all the previous y does not look straightforward.

[3] This assumption is a generalization of Chamberlain random effect model (1980) in

adding and subtracting the term $E[\eta_i|w_i^t]$, we get the following representation for the accident probability:

$$p_{it} = P(y_{it} = 1|w_i^t) = \Phi(\frac{\alpha + \beta y_{it-1} + \sum_{j=2}^{J} \gamma_j k_{jit} + E[\eta_i|w_i^t]}{\sigma_t}).$$

The above equation can be inverted and obtain

$$E[\eta_i|w_i^t] = \Phi^{-1}(p_{it})\sigma_t - \alpha - \beta y_{it-1} - \sum_{j=2}^{J} \gamma_j k_{jit}.$$

By applying the law of iterated expectation we obtain:

$$E[E[\eta_i|w_i^t] - E[\eta_i|w_i^{t-1}]|w_i^{t-1}] = 0.$$

The above equation implies the following moment conditions can be used to estimate the parameters:

$$E[z_{it}(\Phi^{-1}(p_{it})\sigma_t - \Phi^{-1}(p_{it-1})\sigma_{t-1} - \beta \triangle y_{it-1} - \sum_{j=2}^{J} \gamma_j \triangle k_{jit})] = 0 \qquad for\ t = 3...T,$$

where $z_{it} = 1\{w_i^{t-1} = \varphi_k^{t-1}\}$. φ_k^{t-1} is by definition the value that w_i^{t-1} takes.

Identification of α is obtained by the moment conditions $E[\eta_i] = 0$ for $t = 2...T$. In general if we have a time invariant strictly exogenous regressor, its parameter can be identified by the moment condition $E[\eta_i X_i]$ $t = 2...T$. It is important to notice that the moment condition above requires the estimation of p_{it}. The probability p_{it} is estimated non-parametrically using the sample frequencies of y_{it} for all the policy holders with the same history w_i^t.

6 Estimated Results

The estimation method proposed by Arellano and Carrasco is computationally intensive, which necessitated reducing the number of data points. I therefore restrict my analysis to the contracts of new entrants in the year 2000 that remained in the insurance company's portfolio for at least 5 years.

which the unobserved etherogeneity is assumed to depend linearly on the predetermined regressors.

A second simplification I make is to regard claims as a binary variable, because very few policy holders experience more than one claim in each single year. Finally for computational reasons I grouped the number of experience rating classes into 6 categories. The number of experience class indicator in the GMM estimation is hence 5 dummies with as reference group people belonging to lower classes[4].

For purposes of comparison I estimate the model using the following alternative methods: random effect logit estimator (column 1), fixed effect logit estimator (column 2) and Arellano and Carrasco (column 3). Table 5 presents the estimates of the full model including state dependence and the experience class indicators. Table 6 reports the estimates obtained under the different model specifications without state dependence.

The results of the standard random effect probit, which assumes orthogonality between the unobserved heterogeneity and the initial conditions, show that there is negative state dependence but not at a significant level. The parameters of the experience rating classes are estimated to be mostly positive and significant statistically at conventional level, with the exception of the last class. This suggests that there is some dependence between the experience rating class and accident probabilities. However, the magnitude of the coefficients is small and, because spurious positive correlation could inflate the estimates of the experience rating upwards, the use of the random effect probit could lead us to infer that there is some dependence between the experience rating class and accident probabilities, when, in fact, there is none. Column 2 reports the estimates of the fixed effect logit estimator, which allows the unobserved heterogeneity in the population to be correlated. It obtains very different results. Namely, the parameter of the state dependence is negative and significant, indicating that policy-holders having an accident will experience a lower probability of having an accident in the following period. Also all the coefficients of the experience rating scheme result negative, consistent with the fact that people who face a higher marginal increases in the premium will react by exerting more effort and thus having a lower probability of having an accident. However, these estimates could also be biased, as the fixed effect logit model assumes that the experience rating class and state dependence are strictly exogenous when instead

[4]The other 5 groups were constructed as follows: class 1 and 2A were grouped in one class named $k2$, classes 2 through 5 were grouped in $k3$, classes 6 through 9 were grouped in class $k4$, classes 10 through 12 were grouped $k5$ and classes 13 through 18 were grouped in class $k6$.

they are predetermined. This assumption is clearly violated in the model at hand as the experience rating system is function of all the past y. Column 3 reports the estimates of Arellano and Carrasco's estimate, which now allows for a correlated random effect and for predetermined variables. The monotone behavior of the estimates is preserved, but the magnitude of the estimates is very different. This shows that the amount of bias introduced by imposing the strict exogeneity assumption is considerable.

To explore the importance of allowing for state dependence, I estimate the model with and without state dependence. Table 6 reports the estimates without allowing for state dependence. The fixed logit estimator still yields parameter estimates that display a decreasing monotone pattern. Column 3 of table 6 reports the estimates of Arellano and Carrasco estimator without the state dependence. For the most part, there are not significant differences between these estimates and those of the more general model.

To summarize, I find that the estimates of the dependence on experience class do vary by estimator and it is important to use the more general estimator that allows for correlated unobservable heterogeneity and for predetermined variables as implied by the theoretical the model. The estimates based on the Arellano and Carrasco estimator show strong evidence of moral hazard, with a pattern that the probability of an accident is decreasing in the experience class. Individuals who face a higher marginal price of an accident drive more safely, controlling for the unobservable heterogeneity. There is also evidence for state dependence, namely that people who had a recent accident also drive more safely.[5]

7 Sample Selection

One of the potential limitations of my data set is that it pertains only to one insurance company, and it is possible that the occurrence of a claim may not only impact the incentives to follow a precautionary behavior but it may also change the motivation to look for a different insurance company. In an environment where insured may escape the penalties inherent to the experi-

[5] One potential concern is whether the estimates are sensitive to the particular grouping of experience classses. I therefore tried different sets of grouping and the results did not change substantially. Second, the results could depend on the number of groups used. The results did not change substantially if I used a larger set of groups estimated on a smaller size sample (randomly drawn from the original sample).

ence rating system by switching provider it is unclear what bias will result. However, in the Italian market firms share driving records of customers. In particular, when a consumer switches insurance company he needs to provide a document certifying the number of claims experienced in the last 5 years and the so called "ministerial" rating class to which he belongs. The ministerial rating class is composed of 18 classes which coincide with the table of the experience rating class of the company under study beginning from class 1 until class 18 with the exception of class 2A. Therefore, the concern that consumer my escape their claim record is substantially reduced.

Nevertheless, we explore whether the decision to exit the insurance company is related to observable variables in the data, including the response variable (accident claims). Table 7 provides the estimate marginal effects obtained by estimating probit models for the exiting process in each single year of observation. One interesting result is that age and gender have a low impact on the exit probabilities. Another result is that people who experience a claim in a given year have a pretty low increase in the probability of exit. This somewhat suggests that there is very little selection on the dependent variable in the regression therefore lessening concerns about sample selection.

Customers belonging to higher experience classes experience a higher probability of exit especially in the first year. This could indicate that there may an over-representation of the good types in the data. In subsequent years, however, there correlation between exit decisions and experience class becomes smaller. This suggests that probably the high exit that the sample experiences in the first year is due to people with low switching costs departing from the company. To the extent that the estimated model controls for unobservable heterogeneity, selection based on permanent components of accident probabilities would not bias the estimates.

8 Conclusions

In this paper I discussed the importance of monetary costs in determining accident probabilities. Using detailed information about claim histories I am able to examine the empirical importance of moral hazard, adverse selection and state dependence. Both adverse selection and moral hazard are causes of market inefficiencies and can lead to market breakdowns. This paper contributes to the earlier literature by using a rich data set to estimate a more general model than has been previously estimated. With longitudinal data on 300,000 insurance contracts, I can disentangle the separate effects of moral hazard and unobservable heterogeneity on driving behavior. I can also allow for true state dependence, because there is variation in recent accident histories among drivers within the same experience class. The empirical results show a strong evidence that monetary costs are important in determining accident probabilities and therefore that there is moral hazard. The dependence of the estimated results on the estimator used, indicates the importance of employing an estimator whose assumptions accord with the predictions of the model. The Arellano and Carrasco estimator allows for a correlated random effects and for predetermined regressors and yields substantially different results from the standard random effect and fixed effect panel data estimators. The estimates display a decreasing monotone pattern consistent with the observation that consumers that face higher marginal price increases will react exerting more effort and therefore, decreasing accident probabilities. My results also indicate that agents who experience an accident in a given period are less likely to experience another accident in the future period, which suggests that non-monterary costs are also important in determining accident probabilities.

References

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

[2] Abbring J., P.A. Chiappori and T. Zavadil (2007), "Better Safe than Sorry? Ex ante and ex Post Moral Hazard in Dynamic Insurance Data", Mimeo.

[3] Akerlof G. (1970), "The Market for Lemons: Quality Uncertainty and The Market Mechanism", Quarterly Journal of Economics, 84, 488-500.

[4] Anderson, T. and C. Hsiao (1982), "Formulation and Estimation of Dynamic Models Using Panel Data", Journal of Econometrics,18, 67-82.

[5] Arellano M and S. Bond (1991), "Some Tests of specification for Panel Data: Monte Carlo Evidence and an Application to employment Equations", Review of Economic Studies, 58, 277-297.

[6] Arellano M. and R. Carrasco (2003), "Binary Choice Models with predetermined variables", Journal of Econometrics, 115, 125-157

[7] Chamberlain G. (1980), "Analysis of Covariance with Qualitative Data", The Review of Economic Studies, 47, 225-238.

[8] Chamberlain G. (1985), "Heterogenity, Omitted Variable Bias, and Duration Dependence in Longitudinal Analysis of Labor Market Data", edited by J.J. Heckman and B. Singer. Cambridge: Cambridge University Press.

[9] Chamberlain G. (1992), "Comment: Sequential Moment Restrictions in Panel Data", Journal of Business and Economic Statistics, 10, 20-26

[10] Chiappori, P.A. and B. Salanié (2000), "Testing for Asymmetric Information in Insurance Markets", Journal of Political Economy, 108, 56-78.

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

[12] Cooper R. and B. Hayes (1987), "Multiperiod Insurance Contracts", International Journal of Industrial Organization, 5, 211-231.

[13] Dionne G., M. Maurice, J. Pinquet and C. Vanasse (2001) "The Role of Long-Term Contracting with Moral Hazard: Empirical Evidence in Automobile Insurance", Mimeo.

[14] Dionne G., P. Michaud P. and M. Dachour (2007) "Separating Moral Hazard from Adverse Selection and Learning in Automobile insurance: Longitudinal Evidence from France", Mimeo.

[15] Finkelstein A. and K. McGarry (2006) "Multiple Dimension of Private Information: Evidence from Long-Term Insurance Market", American Economic Review 96,4,938-958.

[16] Honore B. and A.C. D 'Addio (2006) "Duration Dependence and Time Varying Variables in Discrete Time duration Models", Mimeo.

[17] Honore B. and E. Kyriziadou E. (2000), "Panel Data Discrete Choice Models with Lagged Dependent Variables", Econometrica, 68, 839-874.

[18] Honore B. and A. Lewbel (2002), "Semiparametric Binary Choice Panel Data Models without Strictly Exogenous regressors", Econometrica, 70, 2053-2063.

[19] Israel M. "Do We Drive More Safely when Accidents are More expensive? Identifying Moral Hazard from Rating Experience Schemes" Mimeo.

[20] Lemaire J. (1995), Bonus Malus Systems in Automobile Insurance, Boston, Kluwer.

[21] Rubinstein A. and M. Yaari (1983), "Repeated Insurance Contracts and Moral Hazard", Journal of Economic Theory, 30, 74-97

[22] Wooldridge J. (2005), "Simple Solutions to the Initial Conditions Problem in Dynamic, Nonlinear Panel Data Models with Unobserved Heterogeneity", Journal of Applied Econometrics, 20, 39-54.

Table 1: Evolution of the experience rating class

Experience rating class	0 claims	1 claim	2 claims	3 claims	4+ claims
1G	1G	1F	1B	2A	4
1F	1G	1E	1A	2	5
1E	1F	1D	1	3	6
1D	1E	1C	2A	4	7
1C	1D	1B	2	5	8
1B	1C	1A	3	6	9
1A	1B	1	4	7	10
1	1A	2A	5	8	11
2A	1	3	6	9	12
2	2A	4	7	10	13
3	2	5	8	11	14
4	3	6	9	12	15
5	4	7	10	13	16
6	5	8	11	14	17
7	6	9	12	15	18
8	7	10	13	16	18
9	8	11	14	17	18
10	9	12	15	18	18
11	10	13	16	18	18
12	11	14	17	18	18
13	12	15	18	18	18
14	13	16	18	18	18
15	14	17	18	18	18
16	15	18	18	18	18
17	16	18	18	18	18
18	17	18	18	18	18

Table 2: Discount on the base premium and marginal price increases

Experience Class	**Discount on the Base premium**	**Percentage increase of the premium**	**Decrease of the premium**	**Marg. Price Increase in Case of a Claim**
1G	0.35	3%	100%	0.01
1F	0.36	3%	97%	0.02
1E	0.37	3%	97%	0.02
1D	0.38	3%	97%	0.02
1C	0.39	5%	95%	0.03
1B	0.41	2%	89%	0.07
1A	0.46	7%	94%	0.08
1	0.49	2%	98%	0.04
2A	0.50	4%	93%	0.08
2	0.54	9%	95%	0.09
3	0.57	9%	97%	0.08
4	0.59	12%	95%	0.09
5	0.62	13%	94%	0.11
6	0.66	12%	95%	0.12
7	0.70	11%	95%	0.12
8	0.74	11%	95%	0.12
9	0.78	13%	95%	0.14
10	0.82	15%	95%	0.16
11	0.88	14%	93%	0.18
12	0.94	22%	94%	0.27
13	1.00	40%	94%	0.46
14	1.15	52%	87%	0.75
15	1.40	50%	82%	0.95
16	1.75	43%	80%	1.1
17	2.10	19%	83%	0.75
18	2.50	0%	84%	0.4

Table 3: Distribution of contracts by experience rating class

Exp. Rat Class	Period 1	Period 2	Period 3	Period 4	Period 5
1G	0	0	0	0	0
1F	0	0	0	0	0
1E	0	0	0	0	3
1D	0	0	0	2	1173
1C	0	0	4	1244	2283
1B	34	2498	3065	2406	634
1A	2611	692	756	596	497
1	698	797	587	414	506
2A	821	525	329	496	292
2	427	320	505	284	289
3	325	490	286	302	426
4	481	273	306	415	377
5	257	306	391	383	389
6	311	387	387	392	337
7	377	401	397	335	244
8	408	406	336	230	242
9	404	328	224	240	286
10	325	206	232	289	532
11	195	217	289	551	69
12	207	296	568	54	76
13	287	583	42	60	101
14	584	33	45	74	12
15	13	10	22	4	3
16	6	4	1	1	1
17	1	0	0	0	0
18	0	0	0	0	0

Table 4: Claim Frequency

Claims	Period 1	Period 2	Period 3	Period 4	Period 5
0	8204 (93.5%)	8251 (94.6%)	8199 (93.4%)	8136 (92.7%)	7999 (91.9%)
1	382 (4.3%)	336 (3.8%)	366 (4.1%)	395 (4.5%)	521 (5.9%)
2+	183 (2.1%)	185 (2.1%)	207 (2.3%)	241 (2.7%)	252 (2.8%)
Total	8772	8772	8772	8772	8772

Note: Percentages in parenthesis.

Table 5: Parameter estimates. Model with state dependence

	RE Estimator	FE Logit	Arellano and Carrasco
Const	-2.74	-	-1.90
	(-75.71)		(-18.75)
$Claim_{t-1}$	-.005	-1.96	-0.37
	(-0.06)	(-21.05)	(-3.91)
k2	0.122	-1.39	-0.45
	(1. 75)	(-10.40)	(-3.59)
k3	0.192	-2.85	-0.71
	(3.19)	(-13.34)	(-3.35)
k4	0.079	-4.58	-1.39
	(1.22)	(-16.65)	(-5.26)
k5	0.158	-5.85	-1.76
	(2.12)	(-17.27)	(-5.52)
k6	-0.084	-7.26	-2.43
	(-0.62)	(-18.33)	(-5.70)

Note: T- and Z- statistics in parenthesis

Table 6: Parameter estimates. Model without state dependence

	RE Estimator	FE Logit	Arellano and Carrasco
Const	-2.76	-	-2.02
	(-79.13)		(-20.41)
k2	0.049	-1.05	-0.47
	(0.79)	(-11.36)	(-3.56)
k3	0.157	-2.03	-0.76
	(2.82)	(-14.27)	(-3.29)
k4	0.148	-3.09	-1.17
	(2.58)	(-16.85)	(-4.05)
k5	0.122	-3.97	-1.79
	(1.76)	(-17.54)	(-5.36)
k6	-0.085	-4.87	-2.38
	(-0.83)	(-18.61)	(-5.30)

Note: T- and Z-Statistics in parenthesis

Table 7 Exit: Marginal Effects

	2001	2002	2003	2004
k2	0.11	0.03	0.02	0.01
	(11.49)	(2.84)	(3.22)	(1.05)
k3	0.10	0.07	0.09	0.04
	(11.51)	(7.89)	(8.84)	(3.58)
k4	0.10	0.08	0.07	0.06
	(11.49)	(8.41)	(6.57)	(4.96)
k5	0.14	0.13	0.08	0.06
	(13.36)	(10.82)	(6.72)	(4.85)
k6	0.18	0.14	0.15	0.12
	(17.36)	(10.57)	(5.29)	(4.04)
Claim	0.06	0.03	0.007	0.007
	(17.34)	(8.61)	(1.51)	(1.09)
Gender	-0.02	-0.04	-0.03	-0.03
	(-4.29)	(-5.62)	(-0.35)	(-4.05)
Age29	0.02	0.04	0.05	-0.002
	(1.78)	(1.94)	(1.65)	(-0.62)
Age35	0.04	0.06	0.06	0.02
	(2.54)	(2.65)	(1.91)	(0.61)
Age40	0.01	0.03	0.04	-0.01
	(0.82)	(1.39)	(1.31)	(-0.36)
Age45	-0.04	0.007	0.01	-0.04
	(-2.99)	(0.36)	(0.37)	(-1.18)
Age50	-0.07	-0.001	-0.005	-0.06
	(-4.66)	(-0.06)	(-0.17)	(-1.76)
Age55	-0.06	-0.03	-0.02	-0.05
	(-4.19)	(-1.83)	(-0.97)	(-1.54)
Age60	-0.05	-0.02	0.004	-0.05
	(-3.22)	(-1.11)	(0.15)	(-1.62)
Age65	-0.05	-0.01	-0.01	-0.07
	(-3.44)	(-0.90)	(-0.48)	(-0.30)
Horse10	-0.03	-0.03	-0.009	-0.007
	(-2.36)	(-2.20)	(-0.43)	(-0.30)
Horse12	-0.10	-0.07	-0.08	-0.07
	(-8.81)	(-5.33)	(-5.40)	(-4.57)
Horse14	-0.11	-0.09	-0.10	-0.08
	(-10.09)	(-7.54)	(-6.98)	(-5.11)
Horse16	-0.09	-0.07	-0.07	-0.07
	(-8.08)	(-5.50)	(-4.69)	(-4.10)
Horse18	-0.09	-0.06	-0.06	-0.05
	(-8.14)	(-4.64)	(-3.92)	(-3.03)
Horse 20	-0.09	-0.07	-0.09	-0.07
	(-5.76)	(-4.24)	(-4.35)	(-3.20)
Horse20+	0.07	-0.02	0.02	0.01
	(2.76)	(-0.83)	(0.54)	(0.35)

Note: Statistics in parenthesis

Fig 1: Life−cycle choice of effort (probability of not having an accident)

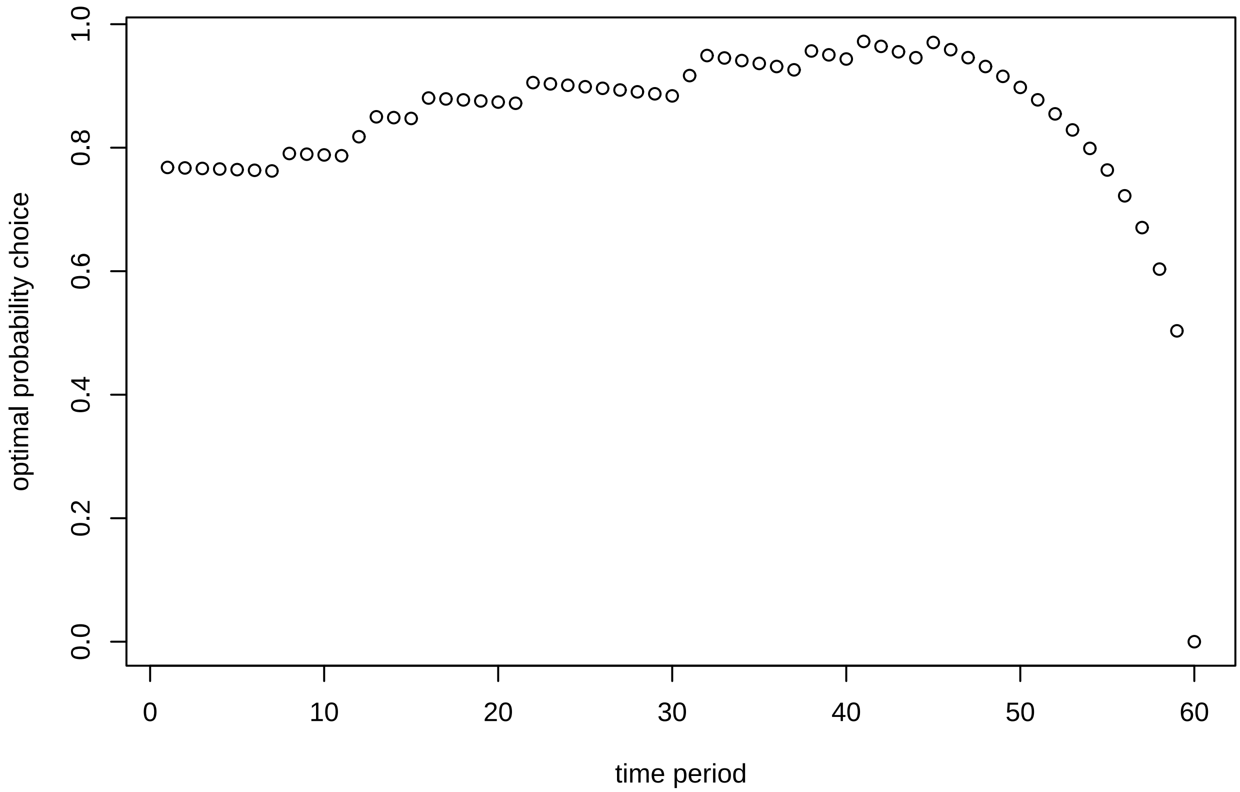

Fig. 2: Relationship between probability and experience class, Time=20