Testing for the Presence of Moral Hazard On Dynamic Insurance Data: Evidence from the Portuguese Car Insurance Industry*

Olivia Ceccarini*
University of Pennsylvania

Nuno Sousa Pereira**
University of Pennsylvania
University of Porto

April 2004

*We are grateful to Eugenio Miravete for helpful comments and suggestions.We would also like to thank Yuichi Kitamura, Mark Pauly and Rafi Rob for invaluable help, Jean Lemaire for the usueful discussion about the Portuguese insurance market, Pierre André Chiappori for the encouragement and Giovanni Millo (from Generali Italy) for conversations about the general functioning of the insurance market. Finally, we would like to thank the Portuguese insurance company that made this paper possible by providing us the dataset. The usual disclaimer applies.

Abstract

This paper tests for the presence of moral hazard in the context of the car insurance market. According to Abbring, Chiappori, and Pinquet (2003), the experience rating system supplies, under mild theoretical assumptions, a straightforward testable implication: in the presence of moral hazard, the occurrence of a claim at any given time is negatively correlated with the occurrence of previous claims. The use of panel data overcomes the well-known problem of disentangling moral hazard from adverse selection. We address the problem of empirically testing for negative state dependence using an extremely rich 4-year panel data set of a Portuguese insurance company. The nature of our data set allows us to employ econometric techniques different from the ones used by Abbring, Chiappori and Pinquet (ibidem), who tested for moral hazard using a 2-year panel data set. The length of our sample enables us to use recently developed econometric procedures for panel data discrete choice models with lagged dependent variables that require at least 3 contractual years for identification purposes.

Preliminary
Please do not Quote

1 Introduction

The importance of adequately distinguishing between adverse selection and moral hazard goes beyond pure scholarly curiosity. Both adverse selection and moral hazard are a consequence of asymmetry of information and can lead to market inefficiencies. However, it is clear that potential inefficiencies arise for different reasons: under adverse selection people with different characteristics choose different contracts, whereas under moral hazard the structure of the contracts causes agents to have different behavior. Moreover, it is widely known that the presence of adverse selection creates an externality problem. Since each insured faces the same unit price, low risk agents subsidize high-risk individuals. In this context, people who create a negative externality are de facto rewarded. Consequently, being able to distinguish between moral hazard and adverse selection would be useful to accurately assess if the presence of asymmetric information causes an undesirable welfare effect at the insured level.

Therefore, it is not surprising that asymmetric information has been at the center of an increasing amount of theoretical research since the seminal papers of Akerlof (1970) and Rothschild and Stiglitz (1976). In contrast to this large body of theoretical work, as Chiappori (2000) pointed out, scholars have only recently addressed the problem of empirically testing the predictions of the theoretical models. The explanation for such a wide temporal lag is two fold. First, it is difficult to establish the presence of asymmetric information, regardless of whether it leads to moral hazard or adverse selection. Second, even after its presence has been established, disentangling those two main effects has proved to be a difficult task.

Interestingly, the scarce empirical studies have shown that most of the propositions widely accepted at the theoretical level are difficult to establish empirically. The issue of empirically testing for the presence of asymmetric information has been addressed by Chiappori and Salanie (2000), who exploit the main theoretical prediction of optimal static contracts. Under adverse selection, high risk agents buy contracts with more coverage than low risk agents. In the presence of moral hazard, agents who buy more insurance will have a smaller incentive to perpetrate a costly action in order to avoid an accident. In either case, and at the empirical level, this translates into assessing whether the choice of the contract is correlated with the accident probability. If, in fact, asymmetric information is present, observationally equivalent agents would choose different contracts. In particular, we should

observe a positive correlation between the accident rate and the quantity of insurance bought. This approach, while offering the advantages of being easily implementable and of not being particularly "data demanding", as it only requires cross-sectional information, is not capable of disentangling moral hazard from adverse selection, since the predictions of the models are the same. Chiappori and Salanie (2000) adopt both parametric and nonparametric techniques to test for the correlation between accident rates and quantity of insurance purchased, but find no evidence of asymmetric information on a sample of young drivers, as well as on a sample of more experience ones.

Cohen (2002), on the other hand, establishes that the presence of asymmetric information cannot be rejected in the Israeli insurance market. The results are mixed and suggest that one important characteristic of insurance contracts is left out of the entire analysis: its real nature. Insurance contracts are long term relationships, and whether there is a full commitment, semi-commitment or no commitment at all should not be ignored. In this sense, Chiappori and Salanie's (ibidem) approach, as the authors acknowledge, fails to take into account this important aspect of the relationship, which probably affects most of their results.

For this exact reason, most of the more recent empirical work on car insurance markets has focused on the dynamic aspects of insurance contracts. At the theoretical level much effort has been devoted to the study of long term contracts. In particular, it has been shown that the properties of optimal contracts are different in the case of adverse selection and moral hazard (Cooper and Hayes (1987), Rubinstein and Yaari(1983)). Focusing on optimal contracts, however, does not usually facilitate trying to perform empirical tests. In fact, optimal contracts usually provide very complicated pricing schemes that are not really observed in actual markets.

Focusing on quasi-natural experiments and simple models of insurance contracting may provide excellent results as in the case of Dionne et al. (2001). The authors build a two period model in which agents are subject to moral hazard and are able to relate the existence of this phenomenon to a change in state regulations. The authors study the introduction of an experience rating system based on demerit points in Quebec and find evidence of the presence of moral hazard in this market. In particular the introduction of experience rating decreased the accident probability by 1%.

The strength of their results rely on two relevant facts. First, in Quebec, because it is the government that sets the bonus-malus scheme, there are no

sample attrition problems. Second, experience rating is based on violations of traffic safety rules, which are a clear measure of moral hazard since they are a good indicator of cautious behavior.

Even though this kind of experiment is rare, the dynamic contracts result is helpful in disentangling moral hazard from adverse selection. This is especially the case if we analyze the induced behavior of the agent taking the exogenously given contract as we see it in actual markets, without having the pretense of assuming any kind of optimality. In particular, as Abbring, Chiappori and Pinquet (2003) show, experience rating schemes based on a previous history of claims prove to be a useful tool in trying to assess the presence of moral hazard. If moral hazard is present, stick and carrot strategies affect the insured's behavior in the following way: the occurrence of a claim affects the entire discounted cost of future claims, making them more expensive. Therefore, the insured will react by exerting more cautious behavior in the following period. On the other hand, the absence of claims affects the whole cost of the premium in the opposite way, making the insured less cautious in the following period. This model, despite its simplicity, provides a powerful prediction that relies on very few assumptions on the primitives. The well- known problem of empirically distinguishing between pure state dependence and spurious state dependence arises when we take the model to the data. In fact, even though the accident probability at any given time should be negatively correlated with the occurrence of a previous claim for each individual, agents differ in their ability to avoid any accidents. More able individuals have a probability of experiencing a claim at any given date and hence this creates a positive spurious state dependence. Therefore, ignoring the source of heterogeneity in the population will cause the estimates to be biased upwards.

Abbring et al. (ibidem) address the issue of testing for negative state dependence using a duration model. The authors are not able to establish the presence of moral hazard in the French insurance market. Our paper differs from the paper by Abbring et al (ibidem) in the methodology used to test for the presence of moral hazard. We adopt the random effects model proposed by Arellano and Carrasco (2003), applying it to an extremely rich longitudinal dataset collected during four years by a main Portuguese insurance company. The detail in the data enables us to correctly identify pure state dependence. Consistent with previous studies on moral hazard we cannot reject the null hypothesis of absence of state dependence in the overall sample. However, further investigation shows that the significance

of the state dependence parameter can be established whenever we limit attention to one particular subset of the sample, precisely the one in which policy-holders do not reach the floor premium in any period of observation. Moreover, a preliminary analysis of the exit of the portfolio indicates that our results could be flawed by the particular nature of the portfolio turnover. People experiencing a claim are more likely to exit the portfolio. We plan to

well.

The insurance premium is composed of two parts at any given time: the base premium, which is computed when the agent becomes the client of the insurance company for the first time, and a bonus malus coefficient also called experience rating. Let P_t be the premium at time t,and let there be two possible states of the world: "accident" and "no accident". The insured is experience rated in the sense that P_{t+1} will depend on the realization of today's state of the world. The premium evolves according to the following dynamic: $P_{t+1} = \gamma P_t$ with $\gamma > 1$ if an accident occurs in $[t, t+1)$, and $P_{t+1} = \delta P_t$ with $\delta < 1$ otherwise. Let $e_t \in [0, +\infty)$ be the level of precautionary activity that the agent can exert at each time t,and $p(e_t)$ the probability of avoiding an accident. Let's assume that $p' > 0$ and $p'' < 0$. The cost of exerting care is given by ve.[1] Under moral hazard the effort level is not observed by the insurer. At each date the insured needs to decide how much effort to exert conditional on the premium dynamic.

It is possible to show that today's premium is a sufficient statistic of the entire past history. Moreover, the today's premium is a function of e_t only. This allows us to write the following Bellman equation:

$$\begin{aligned} V(P) &= \max -ve + (1-p(e))\ u\ (1-P-D) + p(e)\ u\ (1-P) + \\ &\quad + \beta[p(e)V(\delta P) + (1-p(e))V(\gamma P)] \end{aligned} \tag{1}$$

Since V is a contraction mapping, the solution exists and is unique. Moreover, it is possible to show that the value function is decreasing and concave in P. Hence by taking the first order conditions of V with respect to P we have the following identity:

$$-p'(e) = \psi(e) = \frac{v}{u(1-P-D) - u(1-P) + \beta(V(\delta P) - V(\gamma P))} \tag{2}$$

Since V and u are concave we have that if we differentiate ψ with respect to P we obtain:

$$e'(P) > 0 \tag{3}$$

This result provides the following testable implication: the accident probabil-

[1] Both versions allow for a generic cost function with the usual convexity properties. The results, in these cases, hold true for sufficiently small value of the premium.

ity at period $t+1$ is negatively correlated with the occurrence of an accident in the previous period.

This model, as mentioned by Chiappori, represents only a simplified version of reality. For example, it is very common in insurance markets having caps and floors for the premium. As the authors acknowledge the introduction of such kind of dynamics make the model hardly tractable. However, the authors provide some conditions under which the restricted dynamics can be approximated by the unrestricted ones. What the authors basically assert is that for intermediate values of the initial premium (i.e. for premia that are sufficiently far from the cap and the floors) the accident probabilities of the two dynamics differ by an amount of less then ε. As we will see below, this proposition will turn out to be extremely useful in our case.

2.2 The Econometric Model

The purpose of this section is to describe the econometric methodology that we use in order to test for the presence of moral hazard. We will briefly justify the choice of the econometric technique proposed by Arellano and Carrasco (2003) by comparing it to some of the alternative methodologies available. More specifically we can write out the econometric model as follows:

$$\begin{aligned} y_{it} &= 1\{\alpha + \beta z_i + \gamma y_{it-1} + u_{it} + \eta_i \geq 0\}, \\ &\quad \text{with } u_{it} + \eta_i |\ w_i^t \tilde{}\ N\ (E[\eta_i|w_i^t], 1) \end{aligned} \tag{4}$$

where y_{it} is a dichotomous variable taking value 1 if the individual experienced at least one claim in the contractual year t, and zero otherwise. z_i is a set of time invariant dichotomous variables of cardinality k , $w_i^t = \{w_{i1}, w_{i2} .... w_{it}\}$ with $w_{it} = (y_{i,t-1}, z_i)$. The sequence of conditional means $E[\eta_i|w_i^s]$ is left parametrically unrestricted and this is the only non parametric assumption in the model.

It follows that:

$$P(y_{it} = 1|w_i^t) = \Phi(\alpha + \beta z_i + \gamma y_{it-1} + E\ [\eta_i|w_i^t]) \tag{5}$$

Since all regressors are dichotomous, w_i^t has a finite countable support of $2^{t-1}(2^{k+1})$mass points given by φ_j^t

Following Arellano and Carrasco (*ibidem*) let us define:

$$p_{jt} = P(y_{it} = 1|w_i^t = \varphi_j^t) = h_t(w_i^t = \varphi_j^t) \quad j = 1...2^{t-1}(2^{k+1}) \tag{6}$$

Since

$$E[E[\eta_i|w_i^t]] - E[\eta_i|w_i^{t-1}]|w_i^{t-1}] = 0 \tag{7}$$

we have that

$$E[z_{it}(\Phi^{-1}(h_t(w_i^t)) - \Phi^{-1}(h_t(w_i^{t-1})) - \gamma(y_{it-1} - y_{it-2}))] = 0 \quad t = 3,..T \tag{8}$$

where $z_{it} = 1\{w^{t-1} = \varphi_j^{t-1}\}$.

The moment conditions above allows to identify the parameter γ.

Identification of α is guaranteed by the moment condition $E[\eta_i] = 0$ for $t = 2...T$

In order to identify the parameters of the time invariant regressors we made an assumption *à la* Hausman-Taylor. Specifically we assumed that:

$$E[\eta_i|Z_i] = 0 \tag{9}$$

Hence, the time invariant regressors are identified by the following moment conditions:

$$E[Z_i \ \eta_i] = 0, \ t = 2,..T \tag{10}$$

The choice of this estimator over others available is motivated by several considerations. With respect to other random effects estimators (Woodridge (2003)), the one proposed by Arellano and Carrasco offers the advantage of not explicitly modelling the distribution of the unobserved heterogeneity. In this respect, we allow for some form of flexibility. When compared with other fixed effects estimators (Honore and Kyriazidou (2000), Honore and Lewebell (2002)) Arellano and Carrasco's estimator still allows for estimation of partial effects.

3 The Portuguese Insurance Industry

Most insurance companies base their pricing policies on two factors: the base premium and the bonus malus factor. The base premium is usually

evaluated considering the characteristics of the insured, the characteristics of the vehicle, and the chosen amount of coverage. The bonus-malus coefficient is evaluated according to the number of accidents the insured had in the reference period. In particular, the insurance company provides a discount on the future base premium if no accident occurred or otherwise adds an extra charge if the insured had an accident in the reference period.

Base premia usually have the objective of classifying the insured into his respective class of risk, and, hence, it is usually considered as a useful instrument to distinguish high risk individuals from low risk individuals. Experience rating schemes, on the other hand, have multi-fold objectives. As described by Lemaire (1995), they introduce a posteriori measure of the real risk of the insured by linking each single individual to her history of claims. In this respect, they play the same role as the base premium, even though they should be related to each single class so that they can be actuarially fair. The second objective of bonus-malus schemes is to provide incentives for exerting cautious behavior in the presence of potential moral hazard behavior. The third function is to discourage cooperative behavior between policy-holders. Insurance fraud is a disastrous consequence for most of insurance companies, but this is an aspect often neglected by many economists.

We propose to empirically evaluate the potential presence of moral hazard in the Portuguese car insurance market, where insurance companies adopt experience rating schemes in order to price third party liability policies.

In the Portuguese auto insurance market, each insurance company is free to decide on the pricing system they want to adopt, according to their own cost structure and past driving record of their clients. Both base premia and bonus-malus factors are freely set by each insurance company, but their implementation is subject to prior approval by the supervisory authority.

Another institutional feature is that insurance contracts in Portugal are made without commitment. At each renewal date the insured can decide to leave the company and the insurance company can decide whether or not to renew the contract. However, since in Portugal third party liability is mandatory, the insured can appeal to the authority and be reenrolled in insurance under conditions to be determined by the supervisory authority.

Contrary to what happens in most European countries, the Portuguese insurance market is characterized by fierce competition at the bonus-malus factor level. Although segmentation is also present, contrary to what happens in most other European countries very few variables are used in the construction of the base premium. For example, there is no indication of

use of marital status or the occupation of the insured in our records. The determinants most commonly used to set the base premium to be paid are the age, sex and experience of the driver, the age and capacity of the vehicle, the type of coverage, and the magnitude of the deductible. Changes in the premium value can only occur at the end of the contractual year, after notice has been provided by the insurance company, unless otherwise stated in the contractual arrangement.

The bonus-malus factor evolves accordingly to the following dynamic: if the insured experiences a claim during the contractual year the base premium in the following year is increased by 20%. If the insured causes more than one claim the base premium will be increased by 40%. If the insured did not cause any accident the base premium is decreased in the amount of 5%. There are both caps and floors; in particular, the accumulated discount cannot exceed 50% and the accumulated surcharge cannot exceed 250%. An insured that enters the insured company for the first time is given a coefficient of 1 for the first year and then, in the following years, she is experience rated, as all other policy-holders, according to the described dynamics of the evolution of the bonus-malus coefficient. For the time frame of our study, the bonus-malus rule was kept unchanged.

Contracts are experience rated but, if an insured changes company, her previous history is not necessarily available to the new company, unless the insured voluntarily decides to disclosure her previous driving record. Even though insurance companies do not commit explicitly to a particular price or to a particular bonus-malus factor, in some sense, insurance companies commit to the whole bonus-malus strategy.

Table 1 presents some measures of the recent evolution of the Portuguese insurance market. The acute competition mentioned above is visible in both the decreasing number of companies in the market and in the slow growth of the amount of premia received by contract. In 2002, there were 24 insurance companies under Portuguese control operating in the car insurance market, as well as 6 others from other European countries, which represents a decrease of 17% when compared with the year 2000. This decrease is relatively weaker among the Portuguese insurance companies than among the other European insurance companies. The average premium per contract decreased close to 4% in 2000, but it quickly recovered its real value after an increase of 33.7% in 2001. This increase in the premium average value is matched by a significant decrease of 16% in the number of contracts. In 2001, the total number of insurance contracts was equal to 4 815 617, corresponding to a

growth of 2.24%, a very similar growth rate to the one exhibited by the average premium. Although third party liability is mandatory in Portugal, there seems to be strong evidence of high price-elasticity. The decreasing profitability can also be inferred by the higher growth of the amount of euros paid with claims with respect to the amount of euros gathered with the collected premia. While the former faced an annual growth of 10%, the latter exhibited an average growth of 8.25% per year.

4 Data and Descriptive Statistics

Our data base consists of longitudinal data coming from a file of a large, well-established, Portuguese company over a 4 year period. The data includes complete information on the insurer's operations from June 1999 to June 2003, namely details on the characteristics of the contracts (type of coverage, amount of deductible, value of premium, exposition to risk), some characteristics of the individuals that buy the insurance policy (date of birth, gender, date of obtaining driving license, number of claims during the period of analysis, zip code of residence, number of vehicles insured), description of the car (size of engine, model year, value of the car, commercial vehicle or not, number of seats), realization of risks covered (date of claim, date of accident, number of claims, amount of damages paid by the insurance company, whether the insured was at fault or not). We limit our sample to automobile insurance, and we ignore either cars were owned by companies or bought under leasing.

The panel is unbalanced in the sense that not all contracts remain in the insurance company's portfolio for all the observation period. The data contains information on more than 500,000 policyholders that contract with the company during this period of which close to 220,000 remain in the insurance company for the whole observation period.

During the four years under study, the company's number of clients grew at a rate superior to the market's growth rate, which could suggest a more aggressive policy to capture new clients, with the consequent risk of increasing the proportion of high risk individuals in the portfolio. However, Table 2 shows that, during the four years of the analysis, there are no significant changes in the mean characteristics of the company's clients, although some of the more notorious market trends can be observed in the company's portfolio

of clients. First, women are becoming a growing share of the company's clients. Although they are still a minority, representing 15.5% of the total clients, there is an average increase of 1% per year during the time of the analysis. The relevance of this pattern relies on the diverse reaction of men and women to the bonus-malus system, as will be shown in posterior sections. Second, there is a slow but persistent aging of the insured clients. That is visible not only in the mean age of the drivers, but also in their driving experience, with both variables increasing approximately two years during this time frame. Third, there is a lower rate of exits from the company, which leads to a higher mean duration of the contract between the company and its clients.

There is, however, one variable that exhibits a relevant change. The average bonus decreases from 67.888 to 59.702. This is not only a strong evidence of the type of competition that we mentioned in section 3, but also that the clients that exit the company are the ones with higher bonus-malus coefficient, suggesting that there might be a process of risk selection happening as well.

5 Results

5.1 Testing for the Presence of Moral Hazard

In order to implement the method proposed from Arellano and Carrasco we selected the contracts that remained in the insurance company's portfolio for at least 3 years of our time frame and eliminated all the contracts for which some of the variables were missing. We also eliminated from our sample all policy-holders that own more than one car. The motivation for this procedure is that once the insured suffers an accident, all the policies are penalized, regardless of the car involved. This way we avoid a set of inconsistencies regarding the relationship between the evolution of the bonus-malus factor and the number of accidents pertinent to each policy.

A second set of exclusions was implemented in order to make feasible the implementation at computational level. Some insured changed car in the period of observation, and this could be tracked down through several things: the age of the car, type of engines. Even though Arellano and Carrasco' model theoretically allows for such kind of variations, the inclusion of these changes increases the number of possible histories w^t exponentially, making

it impossible to estimate the model due to limited workspace memory. For this reason, we decided to delete all the observations in which there was a change of car. Moreover, we ignored all contracts for which owner and driver do not coincide. These rounds of deletions reduced the number of contracts on which we performed the estimation of the model to 87,914.

A second simplification was regarding our independent variable. Even though the number of chargeable claims can be, in theory, bigger than one, the assumption that the variable in question takes only two values does not seem to substantially affect our results. In fact, on average, 80% of the people with at least one accident in the year had only one claim.

As explanatory variables we used some of the information available on the characteristics of the insured and of the vehicle. As characteristics of the insured we used the dichotomous variable Gender, which equals one if the insured is a female; two dichotomous variables measuring age, Age 1 which is equal to one if the insured was younger than 26 years old when entering insurance, and Age 2 which is equal to one if the insured was between 26 and 56 years old when entering insurance; two dichotomous variables measuring tenure, Tenure 1 which is equal to one for insured who started the contract in 1999 or 2000, and Tenure 2 which is equal to one for insured who started the contract before 1999. As characteristics of the car we use the dichotomous variable Type of car, which is equal to one if the car has up to 1500 cc and 85; two dichotomous variables for power, Power 1, which is equal to one if the car has a cylinder below 1200cc, and Power 2, which is equal to one if the car has a cylinder above 1200 cc and below 1600 cc.

Table 3 presents the GMM estimates based on the entire sample which is composed of 87,914 observations, with t-statistics in parenthesis. It can be seen that the state dependence variable takes negative values, but it is not significant. All the other variables are significant and have the expected sign except gender. Age, as we would predict, has a negative impact on the probability of having an accident. New policy-holders have on average a higher probability of having an accident, more expensive types of car have a higher probability of experiencing a claim. The most striking result is that women have, on average, a higher probability of experiencing a claim.

A further analysis shows, however, that by selecting a sample according to the value the insured had at the initial period, basically deleting the tail of the people who reach the floor in every single period, does not change the value of the parameter in terms of sign. This coefficient remains negative, but becomes significant. The second column of Table 3 presents the results

when we limit our new sample to individuals who did not reach the floor premium in any period of observation. This new sample is reduced to 78,998 contracts.

In this case we can observe that, while all other estimates do not change considerably, the value of the state dependence decreases and the t-statistic indicates a significance at the 5% level(one side test). These results suggest two things. On one hand, the accident probability of the restricted bonus malus dynamic cannot be approximated by the accident probability of the unrestricted one especially for observations in which we observe values very close to the floor at the beginning of the period of observation. A second observation regards the participation constraint. In Chiappori's model the insured must participate by law. Moreover, since in France the bonus-malus system is regulated and the past history of the agent is common knowledge, biases introduced by individuals leaving the insurance company are not very severe. The Portuguese market in this respect is very different. Even though insurance against third party liability is mandatory,competition among firms takes place essentially on the bonus-malus schedule. This implies that problems of sample attrition are severe. Probably, if we could include those individuals who exited the portfolio into our sample, our results would improve.

5.2 Analysis of Exit

One of the potential limitations of any test to the presence of moral hazard is that the occurrence of an accident may not only impact the incentives to follow more precautionary behavior, but it may also change the motivation to look for a different insurance company. In an environment where policy-holders may escape the penalty inherent to the bonus-malus system by switching provider and where each insurance company has its own payment scheme, it is unclear how exactly this fact impacts the results. Therefore, evaluating the process of exit and how it may bias the results of the study and the presence of moral hazard acquires particular interest in this specific market.

In Table 2 we can observe that the percentage of clients that present claims is very steady over time, with the exception of a slight decrease in the last year of observation. This suggests that the relatively small changes in the characteristics of the sample are also reflected by a constant accident rate. If we consider that the bonus-malus coefficient decreases over time, one may infer that there is a constant flow of individuals that subscribe to the

contract each year for the first time, that they are more likely to have an accident, and that they are more likely to leave the company afterwards.

In Table 4, we can observe Probit estimates of the probability of exit for each year of observation, as well as the results for the whole period. The most striking feature of the study of the determinants of exit is that, besides the high significance of all the results, all variables behave as predicted. The occurrence of an accident, a more recent relationship with the company, and a lower driving experience increase the probability that the contract is ended. Furthermore, younger clients, as well as men, are less likely to establish a long term relationship with the company. Finally, there seems to be some support for the hypothesis that individuals with a long term relationship are less likely to leave.

If we consider the results by year, there is no significant change in the magnitude of the coefficients, but the evolution of each one of these factors over time follows different paths. The most obvious changes happen for the variable Accident, which decreases from 0.361 to 0.204, and the variable Female, which evolves from a contribution of -0.558 to a value of -0.229. Both Tenure and Age exhibit a much more steady evolution. Including the bonus-malus coefficient in the regression does not change the results, but confirms the idea that individuals with higher bonuses have higher incentives to switch provider.

6 Conclusion

In this paper, we discussed the limited success of scholars in addressing the relationship between adverse selection and moral hazard. Our methodology and access to data represents an improvement to previous work. First of all, because it relies on a comprehensive four year longitudinal dataset of a Portuguese insurance company, involving more than 500,000 insurance contracts. Second, because the innovative econometric techniques we employ we are capable of distinguishing between spurious state dependence and pure state dependence.

Our preliminary results are not conclusive on the presence of moral hazard. Even though we observe negative state dependence, we only obtain a mild statistical significance when we restrict attention to a particular subsample. A potential explanation for this weak result relies on the bias intro-

duced by individuals exiting the insurance company after presenting a claim. Although we already present results that sustain this argument, we plan to deepen our analysis of this phenomenon through the use of duration models.

References

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

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

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

[4] Chamberlain G. (1985), "Heterogeneity, 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

[5] Chiappori P.A. (2000), "Econometric Models of Insurance Under Asymmetric Information", in G. Dionne, Handbook of Insurance.

[6] Chiappori, P.A. and B. Salanié (2000), "Testing Contract Theory: A Survey of Some Recent Work", World Congress of the Econometric Society.

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

[8] Chiappori P.A. and J. Heckman (2003), "Testing for Moral Hazard on Dynamic Insurance Data:Theory and Economic Tests", Journal of the European Economic Association, forthcoming.

[9] Chiappori P.A., Heckman J. and J. Pinquet J. (2000),"Testing for Moral Hazard on Dynamic Insurance Data", University of Chicago, Mimeo.

[10] Cohen A. (2003), "Asymmetric Information and Learning: Evidence from the Automobile Insurance Market", Mimeo, Harvard University.

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

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

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

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

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

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

[17] Wooldridge J. (2003), "Simple Solutions to the Initial Conditions Problem in Dynamic, Nonlinear Panel Data Models with unobserved heterogeneity", Journal of Applied Econometrics, forthcoming.

Table 1 – Portuguese Insurance Market

	1999	**2000**	**2001**	**2002**
Number of Companies:				
From Portugal	NA	27	25	24
From Other EU countries	NA	9	7	6
Number of Insurance Contracts	4 757 688	5 601 230	4 710 057	4 815 617
Amount of Premia Paid (thousands of €)	1 427 117.80	1 613 657.03	1 815 066.62	1 898 296.48
Claims: Amount Paid (thousands of €)	1 025 790.67	1 231 409.19	1 281 352.18	1 436 908.34

Source: Instituto de Seguros de Portugal (2003)

Table 2 - Sample Means and Standard Deviations

	2000	2001	2002	2003	All years
Number of Observations	241 511	273 136	302 156	304 338	1 121 141
% of clients with claims	7.41	7.99	7.93	6.70	7.50
Female	0.122	0.141	0.155	0.155	0.144
Years as client	5.269	5.083	5.140	5.464	5.242
	(6.427)	(6.520)	(6.567)	(6.590)	(6.533)
Drivers' experience	15.931	16.532	17.134	17.968	16.954
	(10.339)	(10.352)	(10.349)	(10.303)	(10.362)
Drivers' Age	43.032	43.515	44.027	44.875	43.918
	(12.699)	(12.690)	(12.679)	(12.603)	(12.684)
Bonus	67.888	66.196	63.139	59.702	63.973
	(13.404)	(14.680)	(14.460)	(12.876)	(14.220)

Note: Standard deviations are in parentheses below each continuous variable.

Table 3 – Testing for the Presence of Moral Hazard

	Entire Sample	Subsample
Constant	-2.59	-2.51
	(61.93)	(58.3)
State Dependence	-0.019	-0.085
	(0.32)	(1.85)
Gender	0.18	0.16
	(5.3)	(4.5)
Tenure	0.18	0.16
	(7.01)	(5.92)
Age 1	0.35	0.30
	(5.96)	(5.04)
Age2	(0.13)	0.10
	(3.67)	(2.85)
Type of Car	-0.17	-0.157
	(3.56)	(3.19)
Power 1	-0.22	-0.23
	(3.8)	(4.02)
Power 2	-0.15	-0.15
	(-3.08)	(-3.02)
Observations	87,914	78,998

Note: T-Statistics in parenthesis.

Table 4 - Determinants of exit from the insurance company

	2000		2001		2002		2003		All years	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Intercept	-0.580	-1.430	-0.673	-1.741	-0.741	-1.903	-0.826	-1.949	-0.753	-1.753
	(0.012)	(0.021)	(0.011)	(0.020)	(0.011)	(0.018)	(0.012)	(0.020)	(0.006)	(0.010)
Accident	0.361	0.345	0.286	0.251	0.218	0.178	0.204	0.164	0.265	0.233
	(0.011)	(0.011)	(0.010)	(0.010)	(0.010)	(0.010)	(0.011)	(0.011)	(0.005)	(0.005)
Drivers' Age	-0.007	-0.007	-0.006	-0.006	-0.004	-0.003	-0.004	-0.003	-0.005	-0.004
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Tenure	-0.026	-0.017	-0.020	-0.009	-0.016	-0.005	-0.018	-0.010	-0.019	-0.010
	(0.000)	(0.000)	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)
Drivers' experience	-0.002	-0.001	-0.003	-0.001	-0.005	-0.001	-0.003	-0.000	-0.003	-0.001
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Female	-0.558	-0.549	-0.303	-0.305	-0.311	-0.337	-0.229	-0.258	-0.329	-0.342
	(0.012)	(0.012)	(0.010)	(0.001)	(0.009)	(0.009)	(0.009)	(0.009)	(0.004)	(0.005)
Bonus		0.011		0.014		0.016		0.016		0.014
		(0.000)		(0.000)		(0.000)		(0.000)		(0.000)
Year 2000									0.085	-0.023
									(0.004)	(0.004)
Year 2001									0.032	-0.058
									(0.004)	(0.004)
Year 2002									0.062	0.015
									(0.004)	(0.004)
Observations	241 511		273 136		302 156		304 338		1 121 141	
Global Significance	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
Mean Dependent Variable	0.1493		0.1359		0.1396		0.1229		0.1363	

Notes: Dependent variable is one if a client decides to end the insurance contract, and zero otherwise. Year 2003 is the omitted year for regression. All variables are significantly different from zero at the one-percent level.