DO PEOPLE DRIVER SAFER WHEN ACCIDENTS ARE MORE EXPENSIVE: TESTING FOR MORAL HAZARD IN EXPERIENCE RATING SCHEMES*

Tomislav Vukina**

Department of Agricultural and Resource Economics

North Carolina State University

Raleigh, NC 27695-8109

and

Danijel Nestic

Ekonomski Institut Zagreb

Trg. J.F. Kennedy 7

10000 Zagreb, Croatia

---

* The support for this project came through a grant to the University of Delaware from the U.S. State Department's Title VIII Program for 2011. The data set was obtained under the cooperative agreement between Ekonomski institut, Zagreb (The Institute of Economics, Zagreb) and Hrvatski ured za osiguranje (Croatian Insurance Bureau). The authors are thankful to Hrvoje Paukovic and Mario Pavlak for their cooperation in providing the data.

** Corresponding author's coordinates: E-mail: tom_vukina@ncsu.edu; phone: 919-515-5864; fax: 919-515-6268

Do People Drive Safer When Accidents are More Expensive: Testing for Moral Hazard in Experience Rating Schemes

Abstract

Using individual policies and claims data from the Croatian mandatory motor insurance we test the theoretical proposition that under moral hazard, experience rated pricing scheme should generate the negative state dependence in claims, i.e. that drivers should drive more safely after they had an accident. The empirical challenge in these tests is to disentangle the state dependence from unobserved heterogeneity. We propose a simple approach based on the explicit reliance on the cost of future accidents function which is used to filter out the pure incentives effect, whereas the bonus-malus scale is used to control for pure heterogeneity. Our results confirm the existence of negative dependence in claims indicating the presence of significant moral hazard effect.

Keywords: asymmetric information, incentives, insurance

JEL codes: D820, G220

1. Introduction

Multi-period insurance contracts and experience rating schemes have been studied by both economists and the insurance industry (actuaries) since at least early 1980s. From the economics perspective, discounts and penalties are used mainly to mitigate inefficiencies caused by moral hazard (e.g. Rubinstein and Yari, 1983), whereas in the actuarial literature the main purpose of the bonus-malus schemes is to induce people to drive carefully and to better assess the individual risks such that each insured pays the premium which corresponds to his claim frequency (see e.g. Boyer and Dionne, 1989). For both strains of the literature, an important feature of multi-period insurance contracts is that an individual's past experience is a good predictor of the probability of future accidents. The insurer cannot directly observe the self-protection effort of the insured, but the individual's past experience is a good proxy for the individual's risk. Therefore, experience rating schemes enable the insurer to increase the incentives for the insured to exert more care since a good record in the past means a lower premium in the future.

The main objective of this paper is to investigate how people respond to incentives by testing whether they drive more carefully when accidents are more expensive. For this purpose we use a random sample of car insurance policies from all 14 insurance companies that sold mandatory motor third party liability insurance (MTPL) in Croatia in 2009. Croatian system of mandatory motor insurance is based on the experience rating schemes, the so called bonus-malus system. Slightly reformulated, our main objective is to empirically test the theoretical proposition that under moral hazard, bonus-malus pricing system generates the negative state dependence in claims, i.e. that drivers should drive more carefully following the claims. From the policy perspective, the motivation for this paper comes from the necessity to understand the connection between the insurance industry practices and the road transportation safety regulation. There is ample evidence that enforcement of restrictions on alcohol use and speed limits in many countries is rather lax and inefficient and consequently the number of accidents is still quite large. If the results show that drivers are sensitive to bonus-malus insurance schemes, then the road safety regulation should focus on market instruments that are easily implementable and non-arbitrary.

With 12.2 traffic fatalities per 100,000 inhabitants Croatia is comparable to some European countries such Austria (12.2), France (12.9) and Italy (12.2), compares poorly to countries like Sweden (6.7) or United Kingdom (6.1), but compares very favorably to countries like Slovenia (15.8), Poland (16.3) or Portugal (18.1). Croatia made tremendous improvements in terms of reducing the number of traffic related fatalities by more than a half from the peak years of 1979 and 1980 (1,605 and 1,603 respectively) to 548 fatalities in 2009. The number of traffic accidents also dropped precipitously from 92,102 accidents in 2003 and to 50,388 traffic accidents in 2009. This result is even more remarkable taking into consideration the rapid growth in the number of registered motor vehicles throughout the entire transition period. For example, the total number of registered motor vehicles at the end of 2005 was 841,000 whereas that number jumped to 1.4 million at the end of 2009.[1] Such remarkable results are directly related to the construction of the network of new highways that made the travelling in the entire country, but especially the communication between the coast and the hinterland, much more secure. But beside this most obvious explanation, we are also interested in finding out whether more competitive and market driven insurance industry practices in the post-transition period could have contributed to the reduction in traffic accidents.

Empirical tests of the contract theory using insurance data have recently attracted much attention. Several papers test for the existence and estimate the magnitude of asymmetric information effects in competitive insurance markets. Pueltz and Snow (1994), Chiappori and Salanie (2000), Israel (2004), Abbring, Chiappori and Pinquet (ACP) (2003), Chiappori et al. (2006) and Dionne, Michaud and Dahchour (2012), to mention only a few, analyze car-insurance contracts, whereas Cardon and Hendel (2001), Fang, Keane and Silverman (2008), Barros et al. (2008), Liu, Nestic and Vukina (2012) use the health insurance data and Finkelstein and Poterba (2004) focus on annuities.

The approach used in this paper relies on the idea that asymmetric information problems (moral hazard and adverse selection) can be distinguished by analyzing the dynamic aspects of the contractual relationships. The strategy we adopt, based on ACP (2003), takes the existing (and possibly suboptimal) insurance contracts as given and contrasts the behavior implied by the theory to the observed behavior. Specifically, we exploit the fact that under moral hazard,

[1] The numbers are compiled from OECD (2010) and MUP (2009).

experience rating has very interesting implications in that the occurrence of an accident affects the entire schedule of future premiums. This suggests that we can test for moral hazard by testing for such dynamics in the agent's accident process.

The main methodological feature of ACP (2003) is to highlight the distinction between pure heterogeneity and state dependence, the problem that originally appeared in economics in studies of unemployment and labor supply (see Heckman and Borjas 1980). Testing for moral hazard in car insurance with experience rating is similar and boils down to designing a test capable of disentangling the true negative state dependence in accidents (claims) from the unobserved heterogeneity that generates the spurious positive dependence. Using the data from a French insurance company, APC (2003) observe the claim histories for all insurance contracts and are able to model the occurrence of accidents using event-history models. They developed general non-parametric tests which allowed for the non-stationary claim intensity and found no evidence of moral hazard. Their main problem comes from the fact that they observe multiple claims for very few of the many contracts they have in the data which translates into a fairly low precision of their empirical results.

Since the low occurrence of the multiple claims within a contract period (at the end of which the experience rating adjusts) is a common feature of this type of data, we propose a much simpler approach based on the explicit reliance on the cost of future accidents function which is then used to filter out the pure incentives effect, whereas the bonus-malus scale is used to control for unobserved heterogeneity. The fact that the individual's bonus-malus designation is the result of one's historical driving record, and as such endogenous, is dealt with by the use of instruments tied to the nontransparent characteristics of their insurance contracts. After controlling for positive dependence in claims based on unobserved driving ability (proneness to accidents), our results confirm the existence of negative dependence in claims indicating the presence of significant moral hazard effect. The answer to the question posed in the title of this paper is affirmative: yes indeed, people drive more carefully when the accidents are more expensive.

2. Mandatory motor insurance in Croatia

The insurance sector in Croatia is growing but is still somewhat retarded relative to where it should be given Croatia's GDP. According to HANFA (Hrvatska agencija za nadzor financijskih usluga - Croatian Financial Services Supervisory Agency), the relevant supervisory agency responsible for non-bank financial institutions, the Croatian insurance sector settles approximately 607,000 claims every year on 8 million policies (based on 2008 figures) for a population of approximately 4.5 million. There are 25 insurers and 2 reinsurers registered and licensed in Croatia. Among the 25 insurers, 10 have grandfathered composite status such that they underwrite both the life and non-life insurance contracts, 8 are specialized in non-life insurance and 7 in underwriting life insurance contracts. The sector shows typical early transition characteristics for a post-socialist country, i.e., premium revenues are dominated by mandatory motor insurance (MTPL) and motor comprehensive and collision insurance (Casco). In 2008, motor MTPL and Casco insurance combined amounted to 42% of the monetary value of all gross written premiums and the motor MTPL alone amounted to slightly over 30%.

There are 14 insurance companies selling mandatory motor third party liability insurance. The industry is regulated to the extent that the calculation of MTPL premiums is based on the common pricing scheme. The scheme is extremely simple and consists only of two components. The first is the so called "functional premium" and depends on some technical characteristic of the vehicle (power of the engine in KW for cars), the registration region (license plate) and the bonus-malus degree. The second component, the so called “overhead surcharge”, is left to be determined at the discretion of an individual company but must be in the 10-20% range above the functional premium. By default, the duration of the insurance policy is one year, although policies for shorter intervals exist and are sold at a pro-rated basis.[2] The data for the number of insurance policies sold by individual firms and their average premiums for 2009-2010 is presented in Table 1.

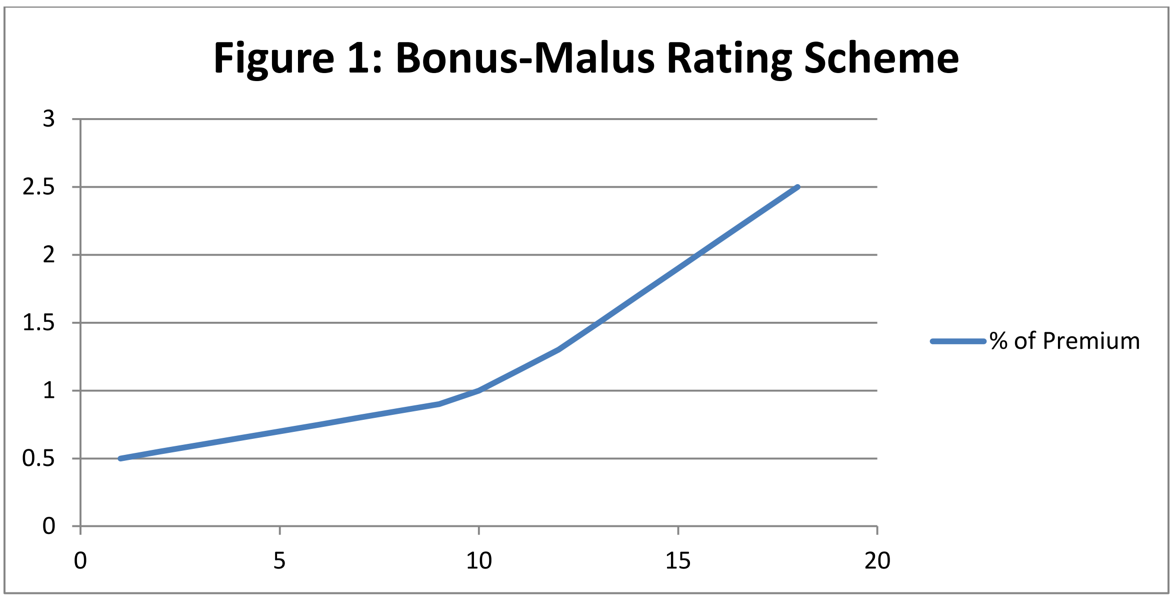
The largest insurance company is the state-owned Croatia Osiguranje with almost 30% market share in 2009 and 28% in 2010. The top-4-firms concentration ratio was 77.8% in 2009 and 76.3% in 2010. The Herfindahl–Hirschman index of 1,880 for 2009 and 1,787 for 2010 indicate moderate concentration. The average annual MTPL premium paid was HRK 1,577 in 2009 and HRK 1,582 in 2010.[3] It is interesting to note that there is some regularity between market shares and prices in the sense that top-3 firms charge below average premiums and bottom-4 firms charge above average premiums but the relationship is not monotonic. The results seem to indicate that firms might have gained market share via price competition but other factors, such as incumbency, reputation, reliability, quality of service, etc., are likely to have played important roles as well.

The Croatian system of mandatory motor insurance is based on the experience rating schemes. The bonus-malus system consists of 18 premium degrees. A new driver who buys the insurance policy for the first time pays the premium calculated at the base level (degree 10). For

[2] The focus of our attention in this study is passenger cars and the description of the functioning of the system pertains strictly to passenger cars although the differences between passenger cars and other vehicles are quite small. For example, the functional premium for trucks is based on weight and for motorcycles on the engine size in cubic centimeters.

[3] In June 2011 the exchange rate for HRK is 5.15 HRK for 1 US$.

each year without a reported accident, for the subsequent policy renewal, the insured receives one degree reduction in premium (bonus), up to the maximum allowable bonus (degree 1). For each reported accident (claim) in the current observation year, for the subsequent year policy renewal the insured receives a three degree increment (malus) up to the maximum allowable malus (degree 18). The bonus-malus scheme with the corresponding percent deductions or increments relative to the base level premium is presented in Table 2 and Figure 1. As seen, the maximum bonus amounts to 50% of the base premium and the maximum malus amounts to 250% of the base premium. It is easy to see that the bonus-malus function has a kink at 10 (base premium) with two almost linear segments where the one on the left (in the direction of lower numbers) has a smaller slope than the one on the right (in the direction of higher numbers). Staying claim-free earns bonuses at a slower rate than having accidents accumulates maluses.



The bonus-malus system generates interesting, possibly perverse, incentives. First, a large number of small, fender-bender type accidents, will go unreported. If the damage to a third party is smaller than the insured's loss due the increased stream of future premium payments, the policy owner (guilty party) will settle the damage with the side payment to the injured party, without ever reporting the accident to the insurance company. Secondly, the acquired experience rating is tied to the owner of the vehicle and to the vehicle itself and not to the driver. For example, the premium for a family car will not be affected by the number and age of the drivers in that family. This means that having an 18-year old child who just started driving will not affect the premium as long as the car is still owned by one of the parents. However, if the child runs into an accident, the parent's premium will be affected by that accident. Thirdly, the acquired bonus status cannot be transferred to the new owner of the same vehicle but the owner can transfer his or her bonus to the newly acquired car when the old one is sold or decommissioned.[4] Finally, switching the insurance company should not change the owner's bonus-malus designation. This means that the bonus-malus degree should "stick" to the agent, in the sense that an agent switching insurers should bring her old coefficient into the new contract. Therefore, in principle, switching insurers could not be explained by an attempt to "escape" the current bonus-malus designation. We revisit this issue later in the paper.

3. Data description

Our data set is extracted from the micro-level database on MTPL insurance for vehicles registered in Croatia, which is held by the Croatian Insurance Bureau (CIB). CIB collects data from all 14 insurance companies that sell the MTPL insurance on the Croatian market. Data exchange between the CIB and the companies is regulated by mutual agreement and is voluntary. Our data set is assembled from two sources – one on the MTPL policies and the other on the claims paid out to injured parties. The combined data set is constructed by linking each claim to the insurance policy of the vehicle/owner identified to be at fault for the reported traffic accident.

[4] Exceptionally, bonus can be transferred together with the vehicle's title to a member of one's immediate family (spouse, parent, child), for example by donation or bequest.

We use a sub-sample of MTPL policies initiated in 2009 and held by private individuals whose family names start with letters R or S and all claims filed against these policies for the entire duration of the 2009 policy coverage.[5] Insurance policies of vehicles owned by businesses, government and similar are not considered. We only consider vehicles owned outright, i.e. vehicles under various leasing arrangements are not considered.

The original dataset contained 149,469 records and needed to be cleaned. First, we excluded all trucks, motorcycles, buses etc. because of the vehicles incomparable technical characteristics (KW for passenger cars, cubic centimeters for motorcycles, metric tons for trucks). Then we excluded all policies with less than one full year of coverage which mainly consisted of insurance premiums for vehicles with temporary license plates. Finally, we excluded all observations with missing or obviously miscued information. The remaining dataset contains 96,433 records (insurance policies) with the summary statistics presented in Table 3.

As seen from the data, the average age of the policyholder is almost 47 and 73 percent of policyholders are men. The average vehicle power is almost 60 kW. The average premium coefficient is 0.55 which means that on average only 55% of the full insurance premium is paid, which is surprisingly close to the maximum allowable bonus of 50%. This is explained by the relatively low rate of reported accidents. Only 3.3 percent of all policies issued in 2009 have reported claims against them.[6] The average insurance premium paid is HRK 1,468, which less than the number reported in Table 1 (HRK 1577). This means that our sample is biased slightly

[5] In Croatia, the insurance policy is related to vehicle and the owner, so one person can have several insurance policies, one for each vehicle. These policies are unrelated, except in the first year when the policy for the second or any subsequent car can be awarded a bonus of 15-25% based on the bonus earned on the previous vehicle. Because of some technical difficulties we are not able to identify all insurance policies belonging to one individual, so we treat all insurance policies as belonging to distinct individuals.

[6] There are 3,041 policies with one reported claim (3.15% of the total), 150 with two claims, 20 policies with three claims and 5 policies with four claims.

downward relative to the entire industry. This makes sense knowing that we deleted all trucks and buses which pay substantially higher insurance premiums than passenger cars.

Table 4 presents the distribution of policies by the premium degree. Astonishingly, 80% of the policies are priced based on the maximum (50%) bonus and over 97% of the policies are priced below the full premium rate. Finally, Table 5 presents the distribution of policies by the registration zone, one of the three critical ingredients determining the functional premium. Zones reflect the level of urbanization and roads congestion such that higher number means higher urbanization and hence higher premium. For example, people whose cars are registered in the registration zone 7 (the highest), which includes the capital city of Zagreb, pay 76% higher base rate than people in zone 1 (the lowest zone).

4. What does the theory predict?

Our theoretical framework relies directly on the model from ACP (2003). The simplified version of this model can be explained as follows. In each period t the agent may, with some probability ($1-p_t$), have an accident. She is fully covered by an insurance contract priced at premium P_t which depends on the past experience. In particular, the evolution of the premium P_t is governed by the bonus-malus system such that $P_{t+1} = aP_t$ if no accident occurred in period t, and $P_{t+1} = bP_t$ if an accident occurred in period t, where $a<1<b$.[7] The no-accident probability p_t is an increasing and concave function of unobservable agent's effort and hence subject to moral hazard. According to bonus-malus scheme, each accident shifts the incentive scheme upward, thus changing the incentives to exert effort. Specifically, the cost of an accident, in terms of higher future premiums, depends on random events (the sequence of accidents) and endogenous decisions (the sequence of efforts).

---

[7] The system actually in place in Croatia is a bit more complicated than this in the sense that the coefficients *a* and *b* can change their values depending on how deep the discount or the penalty is, see Table 2 for details.

After solving the stochastic dynamic control problem, one can show that past experience matters for the current decision only via the current premium and that the optimal effort is increasing in the premium. It follows that the accident probability exhibits negative state-dependence in claims, i.e. that in the absence of an accident, the premium and hence the incentives to exert effort, decreases. The optimal reduction in effort results in the steady increase of the accident probability. On the other hand, the occurrence of the accident generates a discrete jump in the premium which increases the incentives to exert effort and ultimately results in a decreased accident probability. The main testable prediction of this model is that the individual claims intensities decrease with the number of past claims. This implies that we can test for moral hazard by simply testing for negative state (occurrence) dependence in the raw data.

However, at the empirical level, things are substantially more complicated because the individual claim intensities also vary with observed agents' and vehicles' characteristics (age, gender, region, type of car, etc.) and, more importantly, unobservable characteristics (e.g. the innate driving ability). The problem is caused by the fact that this unobserved heterogeneity results in positive state-dependence in the data because those individuals who are bad drivers are more likely to have had accidents in the past and are more likely to have accidents in the future. In other words, to the extent that bad drivers remain "bad" at least in the foreseeable future, we should expect to find a positive correlation between past and future accidents. Therefore, any empirical investigation must address the problem of disentangling the negative state dependence induced by dynamic incentives from the spurious positive dependence generated by unobserved heterogeneity.

What can be learned from the industry data?

A rare and distinct feature of our data set is that it contains a representative sample of insurance policies sold by all 14 insurance companies in Croatia. As mentioned before, the industry is heavily regulated in the sense that all 14 insurance companies must adhere to the same pricing formula. Based on these commonly adopted pricing rules, two tendencies should be immediately obvious. First, to the extent that the functional premium formula is actuarially fair, there should be no reason for the firms to compete to attract better risk pool of drivers because better drivers will, of course, have lower probability of having an accident but they will also pay lower premiums. Secondly, to the extent that the 10-20% overhead surcharge allows firms to earn some normal profit per customer, there should be a clear incentive for the firms to sign-up as many customers as they can regardless of their type.

In light of these two conjectures, it is useful to look at the distribution of contracts (clients) across insurance companies. A simple inspection of the results in Table 6 clearly reveals that this distribution of contracts is not random. First, the market shares differ significantly such that the entire market can be naturally divided into three groups of companies: large, medium and small. This enforces our conjecture that companies compete to gain market share with unequal success.[8] Among the large players (A, D, E), company A (most likely the incumbent – Croatia Osiguranje) stands out with 28.4% of the total issued MTPL policies. The group of medium

[8] During the old socialist regime in former Yugoslavia before 1989, the insurance industry was characterized by the domination of national insurance monopolies. Croatia Osiguranje had a virtual monopoly on the Croatian insurance market, Triglav had a virtual monopoly on the Slovenian market, etc. The presence of the national insurance monopoly from one of the former Yugoslav republics on some other republic's market was generally miniscule. For example, the Slovenian insurer Triglav was present on the Croatian insurance market before 1989 with some tiny, yet precisely unknown, market share. As seen from the largest market share that Croatia Osiguranje still has (Table 1), the incumbency obviously matters. However, some of the newcomers were very successful in spite of the incumbency advantage that Croatia Osiguranje had. Starting from zero, in twenty years or so, they were able to gain a sizeable market share (e.g. Euroherc and Jadransko).

players with market shares ranging from 1% and 6% is the most numerous and appears to be the most homogenous. Finally, three companies (F, I, N) comprise the group of small players with individual market shares of less than 1%. Two of them stand out: company I with zero claims during the 2009 contract year and company N with the highest number of claims in the entire industry (5.1% of its clients had accidents, in comparison to the average accident occurrence for the entire industry of 3.3%).

Next, Table 6 also indicates that the bonus-malus degree is also not evenly distributed among firms. Companies in the large group have somewhat better than average bonus clients than the rest of the industry. The exceptions are company D, from the large group, whose average bonus almost exactly coincides with the industry average bonus and company F, from the small group, which has better than average industry bonus. As mentioned before, the average bonus in the entire industry is very deep and stands at 55% of the regular premium, in comparison to the maximum allowable bonus of 50%.

Finally, the accident occurrence by company shows a rather weak alignment with the average bonus designation by company, the simple correlation coefficient is only 0.386. Based on these numbers we can conclude that there is evidence of some disproportionate distribution of bonus-mauls degrees across firms but the lack of evidence that this distribution coincides with the distribution of actual risk. These findings are also in line with our conjecture that we should not see any systematic distribution of risk across companies as firms have no reason to compete for low risk drivers. This result should hold even if the functional premium is not actuarially fair. If this is the case, then companies with the superior technology to solve the asymmetric information (adverse selection) problems will earn extra-normal profits and possibly expand their market share. However, this does not necessarily mean that their portfolio would consist of disproportionately high percentage of low risk drivers. Instead they would compete for the category of drivers for which the relationship between expected accident occurrence and the regulated functional premium generates the highest expected profits. That could be any category of drivers depending on the bias introduced by the functional premium.

Yet another aspect of competition for clients is worth documenting. As it turns out, switching insurance companies from year to year is fairly common: 14.9% of contracts got switched in 2010 relative to 2009. Recall that switching the insurance company should not change the owner's bonus-malus designation, which means that we should not be able to explain switching insurers by an attempt to "escape" the current bonus-malus designation. Some insights about switching companies can be obtained by the inspection of results in Table 7 where we ran a simple probit regression of the variable indicating the contract switch on the set of drivers', vehicles' and contracts' characteristics. Most of the effects are significant at the standard significance level. For example, having an accident increases the probability of switching the company by 3.6%, being older decreases the probability of changing the company, having a bigger car and having worse driving record (as indicated by higher bonus degree) both increase the probability of changing the company. People having their cars registered in any zone other than zone 1, which is the lowest premium zone, are more likely to change the insurance company than people in zone 1.[9] Finally, people having the insurance policy with any company other than A, are more likely to change the insurer than people who have contracts with company A.[10]

The complete analysis of the insurance industry competition is beyond the scope of this paper, but it is obvious that in this highly regulated market where premiums are allowed to

[9] The zone where the vehicle is registered in most circumstances coincides with the residence of the owner.

[10] Company A also has more clients with the maximum (50%) bonus than any other company.

fluctuate only within a narrow 10-20% range above the predetermined “functional premium”, companies compete for clients with some other non-pricing mechanisms. Three explanations come immediately to mind. First, it is possible that the bonus-malus degree does not “stick” very hard to the agent, i.e. switching insurers can be explained by an attempt to “escape” the current bonus-malus designation, especially in the year after the accident has occurred. By the same token, it is also possible that some companies earn contracts for new vehicles or new drivers by offering bonuses that are larger than they should be if rules were strictly obeyed.

Secondly, it is possible that the services offered by different companies are qualitatively different such that the transaction cost of dealing with an accident in terms of the ease with which the claim is processed varies significantly across companies such that companies offering superior services can out-compete other firms for new clients. On the surface this explanation sounds reasonable, but upon more careful examination of the industry practices appears to be less probable because it is the damaged party who needs to deal with the insurance company to settle the claim and not the owner of the insurance policy who was at fault. Because of this setup, when it comes to MTPL, price should be the most significant, if not the only determinant of the decision from which company to buy the insurance contract.

Finally, the fact that the insurance companies compete on the multiple products markets makes the cross-subsidization of insurance products a viable strategy. For example, all 14 insurance companies, in addition to selling the MTPL policies, also sell motor comprehensive and collision insurance (Casco) and some of them also sell life and homeowners insurance. Therefore the strategy of gaining market share in the MTPL pool is likely to be reinforced by the ability to sell other products to clients whose loyalty has been bought by selling them an inexpensive MTPL policy.

5. Estimation approach

We propose a very simple model which explains the occurrence of accidents A_i^* to an insured i during the time period covered by the contract by his driving ability (proneness to accidents) as approximated by the bonus-malus degree, characteristics of the vehicle, average driving conditions and the cost of potential future accidents:

$$A_i = \beta_0 + \beta_1 B_i + \beta_2 KW_i + \beta_3 Z_i + \beta_4 C_i + \beta_5 C_i^2 + u_i \quad (1)$$

In expression (1), A_i is an indicator variable defined as $A_i = 1[A_i^* > 0]$, $B_i \in [1, 2, \ldots, 18]$ is the insured individual's current bonus-malus degree, KW_i is the power of the vehicle measured in kilowatts, $Z_i \in [1, 2, \ldots, 7]$ is an indicator of the zone where the vehicle is registered, C_i is the cost of future accidents (with the precise definition of this cost given below) and u is the disturbance term.

The main prediction of the model, in line with the moral hazard story, is that the higher the cost (as represented by the schedule of future premiums paid), the lower the probability of having the accident. Ignoring very long term effects of having an accident, we define the relevant cost as the 3-year cost. A 3-year horizon is natural because, as explained before, for each reported accident (claim) in the current observation year, the insured receives a 3 degree increment (malus), up to the maximum allowable malus (degree 18), which will be added to his premium when the policy comes up for renewal in the subsequent period. For each observation year without a reported accident, the insured receives one degree reduction in premium (bonus), up to the maximum allowable bonus (degree 1). This means that each accident, assuming no further accidents occur, will cost the insured 3 points in the first year, 2 points in the second year and 1 point in the third year, before the premium reverts back to the level where it started before the accident. The 3-year cost has the following non-linear structure that depends on the premium *P* and the starting position of the bonus-malus degree *d(k)*:

$$\text{for } d(k=1)\text{: } C_i = [(d_4+d_3+d_2)-3d_1]_i \times P_i$$

$$\text{for } d(k=2)\text{: } C_i = [(d_5+d_4+d_3)-3d_1]_i \times P_i$$

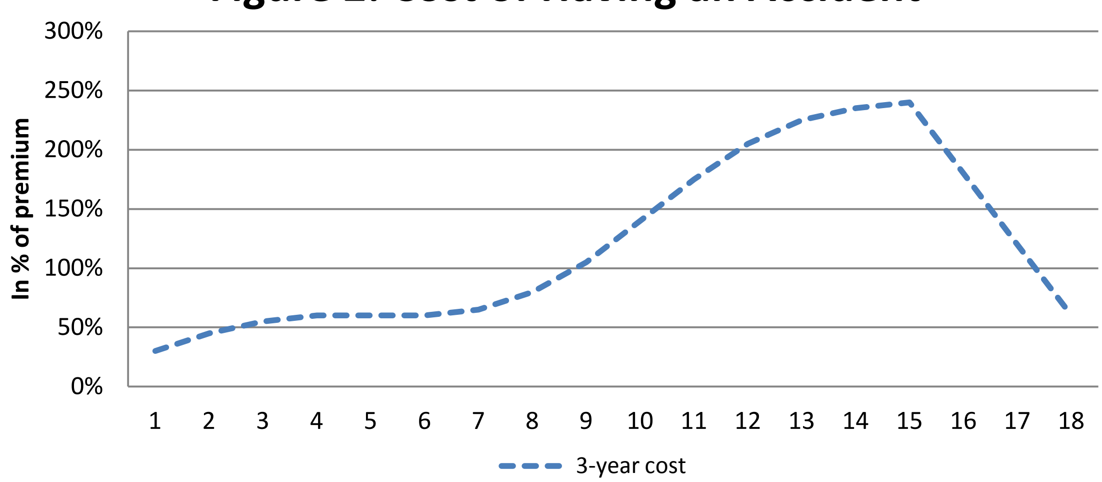
$$\text{for } d(k=3)\text{: } C_i = [(d_6+d_5+d_4)-(d_2+2d_1)]_i \times P_i \qquad (2)$$

$$\text{for } d(k=4,5,\ldots,15\text{: } C_i = [(d_{k+3}+d_{k+2}+d_{k+1})-(d_{k-1}+d_{k-2}+d_{k-3})]_i \times P_i$$

$$\text{for } d(k=16,17,18)\text{: } C_i = [(d_{18}+d_{17}+d_{16})-(d_{k-1}+d_{k-2}+d_{k-3})]_i \times P_i$$

The 3-year cost function in (2) expressed as a percent of any given premium is depicted in Figure 2. As seen, the cost of having an accident as a function of bonus-malus degree is rather shallow up to the degree 7, then increasing up to degree 15 and then decreasing beyond that point. As such it can be represented by some polynomial in *d;* in the empirical part of the paper, we use quadratic because the parameters of the higher order terms ended up being miniscule. In the accident occurrence model, the cost of future accidents function plays the central role of filtering out the pure incentives effect caused by the negative state (occurrence) dependence in claims.

Figure 2: Cost of Having an Accident



The main problem of this approach is how to control for the unobserved heterogeneity in driving abilities. The rationale for the use of multi-period insurance contracts by the insurance industry is the belief that an individual's past experience is a good predictor of the probability of future accidents and hence it can serve as a good measure of one's driving ability. However, the problem caused by including the individual's past experience (bonus-malus designation) as an explanatory variable in the accident equation comes from the fact that it is likely to be correlated with the error term because of the fact that the true driving abilities of the insured are not observable. The standard econometric approach to solving the endogeneity problems involves the use of instruments that must be uncorrelated with u_i and partially correlated with B_i after all other true exogenous variables have been netted out.

Our choice of instruments relies on our in-depth knowledge of the vehicle ownership patterns and driving habits in Croatia and detailed information about the industry structure. We use two types of instruments. The first group includes the observable characteristics of the policy holder, namely age and sex. The reason we believe that these are good instruments follows from the fact that large proportion of the privately owned (not leased) passenger cars in Croatia are actually family vehicles which are typically driven by more than one person. In addition to the person whose name appears on the registration card and the insurance policy (typically a husband), the car is oftentimes driven by wife and older children.[11] Given the fact that equation (1) tries to predict the occurrence of an accident within one period (contract year), the fact that during that period the car could have been driven by multiple drivers, the age and sex of the policy holder are less likely to be correlated with the error term. At the same time since bonus-malus degree is the result of the many years of driving history, it is reasonable to believe that it

[11] The legal driving age in Croatia is 18.

will be correlated with the age and sex of the policy owner. It is easy to think of a situation where the bonus-malus degree could have been determined by the driving history of one of the parents in the family prior to their child becoming old enough to drive. That policy could have earned a maximum bonus based on the perfectly clean driving record of the father and then an 18 year old son started driving affecting the actual probability of having an accident which is not reflected in the bonus-malus degree of the current insurance contract.[12]

The second group of instruments includes the characteristics of the insurance contracts whose detailed variations are not directly observable but can be approximated by insurance companies' fixed effects. In the previous section we have described possible non-transparent competition mechanisms that companies can use to compete for customers and that the result of this competition is reflected in the substantial differences in their MTPL market shares and their portfolios of clients. Some companies are quite large relative to the rest of the industry and some have disproportionately large number of clients with high or maximum bonuses. In addition, switching of contracts between insurance companies is shown to be quite frequent and there are indications that some customers could experience preferential treatment or leniency when buying an insurance contract from a particular company which could improve their bonus-malus rating relative to its exact historical record.[13] As long as there is no systematic matching between companies and driving risk such that having a contract with one insurer indicates a higher risk of

[12]There are many other similar situations that could have produced the same effect. For example, one of the two cars in the family could have been sold because of one spouse losing a job. Now, the remaining car in the family starts being driven by two drivers which non-trivially affects the probability of an accident with no immediate effect on the bonus-malus designation.

[13] The distribution of young drivers in the data set indicates that among 43 18-year old drivers, 33% of them pay less than the full premium (have some bonus) which should not be possible unless the ownership of the vehicle has been transferred from an immediate family member. Given non-trivial transaction costs associated with the ownership transfer, the numbers seem rather high. If we look at the distribution of 21-year olds, among 582 of them, the best that they could do with 3 years of driving with no accidents would be bonus degree 7. As it turns out, 40% of them have deeper bonus than category 7.

having an accident than having a contract with another insurer, company dummies can be considered exogenous.

Finally, notice that the characteristics of the vehicle (power in KW) and the registration zone are not good instruments for the bonus-malus score because people can change cars and transfer their bonus-malus score to another vehicle and they can also change residence which requires them to change the registration plate. Since the bonus-malus score is the result of the long-run driving record, these two indicators do not have to be highly correlated with the historical record but will be very good predictors of the probability of accidents happening in the current period. The speedier the car and the more congested the roads, the higher the chance of having an accident. Therefore, the power of the vehicle and the registration zone are truly exogenous in the equation (1).

The reduced form equation for the endogenous explanatory variable B_i is of the form:

$$B_i = \alpha_0 + \alpha_1 KW_i + \alpha_2 Z_i + \alpha_3 C_i + \alpha_4 C_i^2 + \theta_1 age_i + \theta_2 age_i^2 + \theta_3 M + DF_i \boldsymbol{\theta_4} + \varepsilon_i \qquad (3)$$

where age_i measures the age of the policy holder in years, $M_i = 1[sex = male]$ and $DF_i \boldsymbol{\theta_4}$ are the insurance companies' fixed effects. By definition of the linear projection error, $E(\varepsilon_i) = 0$ and is uncorrelated with all exogenous variables. The econometric model actually estimated is the linear probability model (two-stage least squares). As a robustness check, we also estimated the instrumental variables (IV) probit.

6. Results and discussion

The 2SLS estimation results of equation (1) with the full set of company dummies are presented in the left panel of Table 8a. The corresponding first-stage regression results of equation (3) are presented in Table 8b.[14] As seen from Table 8a, the endogeneity test indicates that the null hypothesis that the bonus-malus variable is exogenous can be rejected, which justifies the use of instrumental variables.[15] Next, the insignificance of some company dummies in the first-stage regression, combined with the F-statistic of less than 10 in the IV estimation, signals the potential problem of weak instruments. As it turns out, the IV redundancy test shows that all 13 companies' dummies are relevant instruments.[16] However, the over-identification test (Hansen J-statistic) forces us to reject the null that the instruments are orthogonal to the error process in the accident equation, ultimately calling the model specification into question.

Given the fact that the large sample bias of the IV estimator increases with the number of instruments, one recommendation when faced with weak instruments is to specify a more parsimonious model; see Baum (2006: pp. 207-211 and references therein). In light of what has been said before about the industry structure, an obvious thing to do was to reduce the number of company dummies by aggregating them into some natural clusters. As seen before, the market really consists of four distinct groups of companies: the incumbent company A, the remaining two large companies (D and E), small players (F, I, N) and the group of 8 medium players. Based on this observation we define: $D_1 = 1[company = A]$, $D_2 = 1[company = D, E]$, $D_3 =$

[14] The estimation was performed with Stata statistical/data analysis package using "ivreg2" command, see Baum, Schaffer and Stillman (2010).

[15] The endogeneity test implemented by "ivreg2" is defined as the difference of two Sargan-Hansen statistics: one for the equation with the smaller set of instruments where the suspect regressor is treated as endogenous and one for the equation with larger set of instruments where the suspect regressor is treated as exogenous. Under conditional homoskedasticity this endogeneity test statistic is equal to Hausman test statistic, see Baum (2006, pp. 211-214.)

[16] The test statistic is a likelihood ratio statistic based on the canonical correlations with and without the instruments being tested. Under the null that the specified instruments are redundant, the statistic is distributed as χ^2 with degrees of freedom equal to the number of endogenous regressors times the number of instruments being tested, see Baum (2006: pp. 207-211).

$1[company = F, I, N]$ and the leftover group containing the medium size companies and use them to replace the full set of companies dummies. The results of a more parsimonious specification of the model (right panels of Tables 8a and 8b) show that the new instruments are all relevant and the F-statistic is now above 10.[17] Also, based on the Hansen J-statistic we cannot reject the null hypothesis that instruments are orthogonal to the error process.

The presented results are very similar between two different specifications. The sign of the bonus-malus effect is positive and the marginal effect is large and significant confirming the positive dependence between current and historical accidents. For example, based on the correct (parsimonious) model specification, increasing the bonus-malus score by 1 degree, increases the probability of having an accident by 15.4%. The power of the vehicle effect is positive and significant indicating that the probability of having an accident increases as the power of the car's engine increases but the magnitude of the effect is quite small: an increase in the engine power by 1 KW increases the probability of an accident by 0.4%. All registration zones' fixed effects relative to the left-out zone 1 (least expensive insurance premium) are also positive and significant in both models. For example, based on the same model specification, having a car registered in Zone 7 where one pays 76% higher base premium, increases the probability of having an accident by 21.5%. This shows that the premium pricing formula used by the industry is on the right track: charging higher premiums for higher powered vehicles and higher traffic density zones seems to be justified because both of those factors are clearly associated with higher probability of accident.

[17] Further reduction in the number of company dummies by dividing the entire market only into 3 (company A, other large companies, and all others) or even 2 groups (company A and all others) further improves the statistical results, although qualitatively, the results do not change. Additional tests for the redundancy of age, gender and company A, individually or in any combination, always reject the null that instruments are redundant at 1% significance level. All those results are available from authors upon request.

The most important result of the paper is revealed in the estimated coefficients of the cost function which is approximated with the second degree polynomial. As seen, both linear and quadratic coefficients are significant and the total effect (the derivative) evaluated at the mean full (degree 10) premium is negative indicating the negative state dependence in claims or the presence of the significant moral hazard effect. Again, using the right panel estimates from Table 8a, measured at the maximum allowable discount (degree 1), the results show that increasing the 3-year cost of having an accident by 100 HRK (approximately US$20) decreases the probability of having an accident by 7%. Interestingly, the effect monotonically decreases as one slides towards lower bonuses such that at the full average premium (degree 10), the increase in 3-year cost by 100 HRK would cause a decrease in probability of having an accident by only 2.8%.

Finally, it is rather interesting to note that for the malus degree 12 (130% of the base premium) and higher, the effect turns positive, i.e. the probability of having an accident increases as drivers sink deeper and deeper into the malus territory. However, one has to be cautious and not overemphasize this result because, as seen from Table 4, the number of drivers in these high malus categories drops precipitously, such that we have only 4 drivers with the malus degree 16, no drivers with degree 17 and 1 driver with degree 18.

In addition to the linear probability model, we also estimated the instrumental variable probit. The main results are presented in Table 9. Same as before, the left panel of the table contains the results of the full companies' fixed effects model, whereas the right panel presents the parsimonious model where the market was segmented into 4 groups of companies (3 dummy variables).[18] As seen from the table, all results are qualitatively identical to those obtained with

[18] The estimation was performed using Stata "ivprobit" command with Newey's two-step estimator. In the probit estimation the full set of company instruments contains only 12 dummy variables (and not 13 like in the linear

the linear probability model, so all our conclusions are reinforced by an alternative estimation approach.

7. Conclusions

In the recent years the empirical tests of contract theory using insurance data have been quite numerous. In most cases researchers managed to assemble cross-sectional data sets, while panel data sets were much less frequently available. In all instances, however, the data came from only one insurance company. In the panel data analyses this presented a challenge of dealing with attrition of clients because, invariably, some people changed the insurance company from year to year, thereby dropping from the available data sets. Our data set is unique in the sense that we have a snapshot of the entire industry. Our data is randomly drawn from the population of all mandatory motor third party liability insurance contracts sold in 2009 by all 14 insurance companies that span the Croatian market. Because the coverage of almost all contracts issued in 2009 expired in 2010, the claims data stretches well into 2010. The same group of people was then followed into 2010 to see which one of them renewed their contracts with the same company and who switched to another insurer. However, we don’t have claims filed for policies issued in 2010 because the period covered by contracts issued in 2010 carried over into 2011 for which no data is available.

By being able to identify the insurance company that issued a contract we found that some companies have substantially larger market shares than others, that motorists are not randomly assigned across companies, and that the probability of changing the insurance company depends

model) plus the left-out company A. This is because company I has zero reported claims and hence its dummy variable would perfectly predict zero claims. To solve this numerical problem, we introduced new (joint) dummy variable for the two smallest companies I and F.

on one's current insurer. These results were critical for model identification. The central problem of the empirical analysis in this type of setting is how to disentangle the pure incentives effect associated with moral hazard from the selection effect caused by unobserved heterogeneity. Under moral hazard, experience rating implies negative state dependence of individual claim intensities which is confounded with the positive selection effect because a driver with large number of past accidents is likely to have large number of accidents in the future. We proposed a simple approach based on the explicit reliance on the cost of future accidents function which is used to filter out the pure incentives effect. To control for pure heterogeneity we used the bonus-malus scale where the company dummies and the gender and age of the insurance policy holders served as instruments. Our results confirm the existence of negative dependence in claims indicating the presence of significant moral hazard effect.

The obtained results could have an important policy implication. For policy purposes it is important to understand the connection between the insurance industry practices and the road transportation safety regulation. There is ample anecdotal evidence that enforcement of restrictions on alcohol use and speed limits in many East European countries are rather lax and inefficient and consequently the number of accidents and violations is still relatively large. Because our results show that drivers respond to incentives created by the experience rating schemes, the focus of road safety regulation should switch from instruments that critically rely on the efficient and non-corrupt traffic police to instruments that are easily implementable and completely non-arbitrary and at the same time could improve the profitability of the insurance sector.

References:

Abbring, J.H., P.-A. Chiappori and J. Pinquet (2003). “Moral Hazard and Dynamic Insurance Data.” *Journal of the European Economic Association* 1 (June): 767-820.

Barros P.P., M. Machado, and A. Sanz-de-Galdeano (2008). “Moral Hazard and the Demand for Health Services: A Matching Estimator Approach.” *J. of Health Economics*, 27 (4):1006-25.

Baum, C.F. (2006). “An Introduction to Modern Econometrics Using Stata.” Stata Press, College Station, Texas.

Baum, C.F., M.E. Schaffer, and S. Stillman (2010). “Ivreg2: Stata module foe extended instrumental variables/2SLS, GMM and AC/HAC and k-class regression. http://ideas.repec.org/c/boc/bocode/s425401.html

Boyer, M. and G. Dionne (1989). “An Empirical Analysis of Moral Hazard and Experience Rating.” *The Review of Economics and Statistics*, 71 (1): 128-134.

Cardon, J. H. and I. Hendel (2001). “Asymmetric Information in Health Insurance: Evidence from the National Medical Expenditure Survey.” *RAND Journal of Economics* 32 (3): 408-427.

Chiappori, P.-A., B. Jullien, B. Salanie, and F. Salanie (2006). “Asymmetric Information in Insurance: General Testable Implications.” *Rand Journal of Economics* 37 (Winter): 783-798.

Chiappori, P.-A. and B. Salanie (2000). “Testing for Asymmetric Information in Insurance Markets.” *Journal of Political Economy* 108 (1): 56-78.

Dionne, G., P.-C. Michaud and M. Dahchour (2012). “Separating Moral Hazard from Adverse Selection and Learning in Automobile Insurance: Longitudinal Evidence from France.” *Journal of the European Economic Association*, 10: forthcoming.

Fang, H., M.P. Keane, and D. Silverman (2008). “Sources of Advantageous Selection: Evidence from the Medigap Insurance Market.” *Journal of Political Economy* 116 (2): 303-350.

Finkelstein, A. and J. Poterba (2004) "Adverse Selection in Insurance Markets: Policyholder Evidence from the U.K. Annuity Market," *Journal of Political Economy* 112 (February):183-208.

Heckman, J. J. and G.J. Borjas (1980): “Does Unemployment Cause Future Unemployment? Definitions, Questions and Answers from a Continuous Time Model of Heterogeneity and Time Dependence.” *Economica* 47: 247-283.

Israel, M. (2004). “Do We Drive More Safely When Accidents are More Expensive? Identifying Moral Hazard from Experience Rating Schemes.” The Center for the Study of Industrial Organization at Northwestern University, Working paper #0043.

Liu, X., D. Nestic and T. Vukina (2012). “Estimating Adverse Selection and Moral Hazard Effects with Hospital Invoices Data in a Government-Controlled Health Care System.” *Health Economics*, Vol. 21 (August 2012): 883-901.

MUP (2009). Republika Hrvatska, Ministarstvo unutarnbjih poslova. “Bilten o sigurnosti cestovnog prometa 2009.” ISSN 1331-2863, Zagreb 2010.

OECD (2010). International Transport Forum. IRTAD Road Safety 2010 Annual Report. www.irtad.net

Pueltz, R. and A. Snow (1994). “Evidence on Adverse Selection: Equilibrium Signaling and Cross-Subsidization in the Insurance Market.” *Journal of Political Economy* 102 (April): 236-57.

Rubinstein, A. and M.E. Yaari (1983). “Repeated Insurance Contracts and Moral Hazard.” *Journal of Economic Theory* 30 (June): 79-97.

Table 1: Industry Structure & Prices

Company	Contracts Sold		Market Share		Avg. Premium (in HRK)		Over/under Pricing	
	2009	2010	2009	2010	2009	2010	2009	2010
Croatia	579,607	545,182	29.6%	28.3%	1,443	1,446	-134	-136
Euroherc	484,007	467,234	24.7%	24.3%	1,398	1,415	-179	-167
Jadransko	326,135	318,709	16.7%	16.5%	1,414	1,420	-163	-162
Allianz	132,952	138,272	6.8%	7.2%	1,709	1,701	132	119
Kvarner	87,213	81,957	4.5%	4.3%	1,590	1,576	13	-6
Triglav	83,526	79,061	4.3%	4.1%	1,490	1,552	-87	-30
Basler	60,412	66,881	3.1%	3.5%	1,531	1,532	-46	-50
HOK	55,424	67,519	2.8%	3.5%	1,618	1,641	41	59
Grawe	38,923	37,712	2.0%	2.0%	1,413	1,414	-164	-168
Generali	37,897	42,511	1.9%	2.2%	1,619	1,644	42	62
Uniqa	23,218	25,936	1.2%	1.3%	1,847	1,817	270	235
Velebit	17,805	22,270	0.9%	1.2%	1,731	1,720	154	138
Sunce	15,012	17,434	0.8%	0.9%	1,639	1,623	62	41
Helios	14,658	15,966	0.7%	0.8%	1,634	1,641	57	59
TOTAL	1,956,789	1,926,644	100.0%	100.0%	1,577	1,582	0	0

Source: Contracts and premiums data from the Croatian Insurance Bureau (CIB). Reprinted from Poslovni Dnevnik, 24-25. June, 2011, p.16.

Table 2: Bonus-malus premium degree schedule

Premium Degree	% of Premium
1	0.5
2	0.55
3	0.6
4	0.65
5	0.7
6	0.75
7	0.8
8	0.85
9	0.9
10	1
11	1.15
12	1.3
13	1.5
14	1.7
15	1.9
16	2.1
17	2.3
18	2.5

Table 3: Data summary statistics

	2009	
Variable	Mean	STD
Age of policyholder (years)	46.90	14.06
Gender (M=1, F=0)	0.729	0.445
Premium coefficient (100% = 1)	0.550	0.124
Accident in 2009 (YES=1, NO=0)	0.033	0.180
Vehicle power (in kW)	59.70	20.33
Premium paid (in HRK)	1468.41	504.52
Number of observations (insurance policies)	96,433	

Table 4: Distribution of policies by the premium degree

Premium Degree	Proportion	
	Percent	Cumulative
1	80.05	80.05
2	2.26	82.31
3	3.07	85.38
4	3.43	88.81
5	1.28	90.09
6	1.60	91.69
7	1.19	92.88
8	2.20	95.08
9	2.29	97.38
10	2.38	99.75
11	0.08	99.84
12	0.11	99.94
13	0.04	99.98
14	0.01	99.99
15	0.01	99.99
16	0.00	100.00
17	0.00	100.00
18	0.00	100.00

Table 5: Distribution of the policies by registration zone

Registration zone	Premium coefficient	Percent share of the total
1	1.00	1.08
2	1.10	5.20
3	1.21	8.79
4	1.33	41.24
5	1.46	12.01
6	1.60	3.40
7	1.76	28.29

Table 6: Distribution of MTPL insurance contracts across insurance companies

Company	**Market Share**	**Average Bonus**	**Policies with Claims**
A	28.4%	54.2%	2.7%
B	2.2%	57.1%	3.5%
C	6.1%	55.8%	3.7%
D	18.4%	54.8%	3.6%
E	27.2%	55.1%	3.4%
F	0.7%	54.9%	3.6%
G	3.2%	55.9%	3.3%
H	2.4%	55.3%	4.1%
I	0.7%	55.1%	0.0%
J	4.4%	56.4%	3.8%
K	1.2%	55.8%	3.5%
L	2.4%	55.6%	4.5%
M	2.0%	56.3%	3.5%
N	0.8%	56.8%	5.1%
Industry Total	100.0%	55.0%	3.3%

Note: The insurance company codes A through N have been randomly assigned and cannot be used to precisely identify companies in Table 1, although some guessing is certainly possible.

Table 7: Switching contracts between insurance companies (Probit regression)

Dependent variable: switch (yes/no)	Marginal effects	*z*-statistic	P>|*z*|
Accident	0.036	6.180	0.000
Age	-0.004	-9.050	0.000
Age Sq	2.6e-05	5.360	0.000
Male	0.002	0.780	0.434
Power	2.3e-04	4.540	0.000
Bonus-malus	0.001	2.060	0.040
Registration zone (Reference zone 1)			
2	0.050	3.180	0.001
3	0.041	2.730	0.006
4	0.050	3.630	0.000
5	0.045	3.020	0.003
6	0.077	4.570	0.000
7	0.058	4.040	0.000
Company (Reference company A)			
B	0.516	47.670	0.000
C	0.330	44.780	0.000
D	0.253	48.530	0.000
E	0.209	45.850	0.000
F	0.494	27.430	0.000
G	0.428	45.900	0.000
H	0.325	31.050	0.000
I	0.314	17.240	0.000
J	0.387	46.170	0.000
K	0.429	31.400	0.000
L	0.426	40.960	0.000
M	0.399	35.740	0.000
N	0.424	25.370	0.000
No. of observations		96,433	
Pseudo R^2		0.097	

Table 8a: Testing negative dependence in claims: Linear probability model (2SLS)

	Full set of company			*Company group dummies*		
Dependent variable: Claim	Coefficient	Robust		Coefficient	Robust	
Bonus	0.096	0.016	***	0.154	0.022	***
3-year cost	- 4.6e-04	8.2e-05	***	-7.6e-4	1.1e-04	***
3-year cost squared	4.01e-08	7.4e-09	***	6.5e-08	1.0e-08	***
Power	0.002	3.8e-04	***	0.004	5.1e-04	***
Regist. zone (Ref: Zone 1)						
2	0.020	0.006	***	0.031	0.008	***
3	0.041	0.008	***	0.061	0.010	***
4	0.066	0.011	***	0.102	0.015	***
5	0.093	0.014	***	0.141	0.019	***
6	0.110	0.018	***	0.171	0.024	***
7	0.138	0.022	***	0.215	0.030	***
Constant	-0.114	0.022	***	-0.192	0.029	***
No. of observations	96433			96433		
F-statistic	8.93			11.87		
No. of over-identifying restrictions	15			5		
	χ^2	P-value		χ^2	P-value	
Endogeneity test [(a)]	19.695	0.000		66.533	0.000	
IV redundancy test [(b)]	169.905	0.000		54.621	0.000	
Hansen’s J-test	233.919	0.000		8.152	0.148	

Notes: *, ** and *** indicate statistical significance at the 10%, 5% and 1%, respectively. (a) See footnote 16. (b) See footnote 17.

Table 8b: Linear probability model (2SLS) – first stage results

Dependent variable:	*Full set of comp. dummies*			*Company group dummies*		
Bonus/Malus	Coefficient	Robust S.E.		Coefficient	Robust S.E.	
Age	-0.006	0.001	***	-0.006	0.001	***
Age squared	5.4e-05	1.0e-05	***	5.3e-05	1.0e-05	***
Male	-0.037	0.005	***	-0.036	0.005	***
Company (Ref: Comp. A)						
B	-0.020	0.016				
C	-0.014	0.009				
D	0.016	0.006	**			
E	0.049	0.006	***			
F	0.002	0.023	**			
G	0.069	0.014	***			
H	0.021	0.015				
I	0.069	0.022	***			
J	0.086	0.011	***			
K	0.025	0.018				
L	0.044	0.014	***			
M	0.012	0.016				
N	0.008	0.024				
Co. groups (Ref: Medium)						
Group 1 (incumbent)				-0.029	0.006	***
Group 2 (large)				0.007	0.005	
Group 3 (small)				-0.003	0.014	
3-year cost	0.005	7.7e-05	***	0.005	7.7e-05	***
3-year cost squared	-4.2e-07	1.9e-08	***	-4.2e-07	1.9e-08	***
Power	-0.023	3.1e-04	***	-0.023	3.1e-04	***
Reg. zone (Ref: Zone 1)						
2	-0.192	0.028	***	-0.192	0.028	***
3	-0.347	0.028	***	-0.343	0.028	***
4	-0.608	0.026	***	-0.607	0.026	***
5	-0.819	0.028	***	-0.818	0.028	***
6	-1.046	0.029	***	-1.044	0.029	***
7	-1.314	0.028	***	-1.315	0.028	***
Constant	1.509	0.044	***	1.534	0.043	***
No. of observations	96433			96433		
Uncentered R^2	0.9580			0.9580		

Notes: *, ** and *** indicate statistical significance at the 10%, 5% and 1%, respectively.

Table 9: Testing negative dependence in claims: Instrumental variable probit

Dependent variable: Claim (Yes/No)	*Full set of company dummies*			*Company group dummies*		
	Coefficient	S.E.		Coefficient	S.E.	
Bonus	1.358	0.211	***	2.055	0.275	***
3-year cost	-0.007	0.001	***	-0.010	0.001	***
3-year cost squared	5.69e-07	9.0e-08	***	8.66e-07	1.2e-07	***
Power	0.032	0.005	***	0.048	0.006	***
Reg. zone (Ref: Zone 1)						
2	0.307	0.109	***	0.439	0.118	***
3	0.616	0.121	***	0.855	0.139	***
4	0.971	0.157	***	1.391	0.193	***
5	1.347	0.196	***	1.913	0.245	***
6	1.596	0.243	***	2.322	0.307	***
7	1.989	0.292	***	2.902	0.375	***
Constant	-3.950	0.295	***	-4.875	0.379	***
No. of over-identifying restrictions	14			4		
No. of observations	96433			96433		
Test of exogeneity: Wald χ^2	46.75	0.000		71.680	0.000	
Overidentification test:						
Amemiya-Lee-Newey min. χ^2	52.983	0.000		8.548	0.129	

Notes: *, ** and *** indicate statistical significance at the 10%, 5% and 1%, respectively.