Testing for Asymmetric Information in Tunisian Automobile Insurance Market

Imen Karaa

*Economic, Management and computer Sciences Doctorate School of FSEG Sfax,
University of Sfax,Route de l'Aéroport Km 4 Sfax 3018 Tunisia
Email:karaaa.imen@gmail.com*

Noureddine Benlagha*

*Department of Finance and Investment, College of Economics and Administrative Sciences,
Al Imam Mohammad Ibn Saud Islamic University (IMSIU), P.O. Box 5701, Riyadh, Saudi Arabia
Email: blnouri2002@yahoo.fr*

Doi:10.5901/mjss.2015.v6n3s1p455

Abstract

Asymmetric information has become one of the most influential theoretical concepts in explaining competitive insurance market. In contract theory, asymmetric information is the situation when one party of an economic transaction possesses information that is not available to the other contractual party.In this perspective, this paper aims to investigate the presence of asymmetric information in Tunisian automobile insurance market. This investigation is made using an individual data set during the year 2009. For this purpose, we employ two empirical approaches; the bivariate probit model and the two stages approach of Richaudeau (1999). Hence, results show a strong evidence for the presence of asymmetric information for the begging drivers. Inversely, for the experienced drivers, no evidence of asymmetric information have been detected.

Keywords: *Asymmetric Information, Tunisian market, automobile insurance, Negative Binomial model, Bivariate Probit model.*

1. Introduction

Asymmetric information has become one of the most influential theoretical concepts in explaining competitive insurance market. In contract theory, asymmetric information is the situation when one party of an economic transaction possesses information that is not available to the other contractual party.

In fact, Asymmetric information arises when an informational gap exists between an insurer and an insured. In such situation, insureds are more informed about their own characteristics or actions than their insurer. Potentially, this could be a harmful situation because the insured can take advantage of the insurer's lack of knowledge. In a nutshell, asymmetric information can leads to a problematic because it creates opportunities for lying and cheating by insured, which by the way can induce market inefficiency or even failure (Arrow 1963).

In this field, a number of empirical studies propose to measure the presence of residual asymmetric information problems in several automobile insurance markets. One can refers to Dahlby (1983/1992) for Canada, Puelz and Snow (1994) for USA, Richaudeau (1999) for France, Saito (2006) for Japan, Kim et al. (2009) for Korea, Englund (2010) for Denmark, Shi and Valdez (2011) for Singapore, and Li et al. (2013) for Taiwan case. Besides a recent review article by Dionne et al. (2013) which summarized the theoretical and the empirical literature concerning asymmetric information in road safety and automobile insurance market.

In the Canadian automobile insurance market, two initial studies (Dahlby, 1983 and 1992) have estimated the sign of the correlation between the level of insurance coverage and ex-post realizations of risk. Puelz and Snow (1994) found positive coverage-risk correlation using individual data obtained from a Georgia insurer.

However, several empirical studies show no evidence for asymmetric information. Using a French data set, Richaudeau (1999) as well as Chiappori and Salanié (2000) find no systematic relationship between coverage and risk. In addition, Dionne et al. (2001) reach a similar conclusion for the Canadian automobile insurance market: those who are more likely to submit claims do not buy more insurance. Moreover, these researchers criticized Puelz and Snow (1994) for failing to consider nonlinear effects.

On the other hand, Cohen (2005) did not find any correlation for beginning drivers[1], she founds a strong correlation for experienced drivers. Although, the study suggested the importance of "learning" in this market.

Saito (2006) used an individual data from the Japanese auto insurance market, where insurance rates are strictly regulated. First, he found a positive but not statistically significant correlation between the crash risk and the purchase of own-vehicle coverage, even under the control of all variables observed by the insurer. Second, he concluded that there was a significant and negative correlation between coverage and risk when they are defined as the purchase of a zero-deductible policy and a crash risk.

Recent empirical studies in the competitive insurance market show a presence of asymmetric information in these markets. To take some of the outstanding studies, Grun-Rehomme and Benlagha (2007) examined the choice between basic third party coverage and deductible variable for beginning and experienced drivers. Their study used 4 guarantees' types concerning the year 2004, this was made after the new circulation security measures' implementation and the decrease of mortals' accidents road. Using a bivariate ordered model, the researchers find a significant positive correlation risk-coverage for beginning drivers which choice the third party insurance plus deductible variable for French data.

We can notice that most of the previous empirical focuses on the coverage-risk correlation. This paper contributes to the literature by reexamining whether such coverage-risk correlation exists.

The aim of this paper is to investigate on residual asymmetric information in the Tunisian automobile insurance market and to find ways to improve and reduce the negative impact of asymmetric information on resource allocation. Hence, the objectives of this research proposal are to test the presence of the adverse selection in the Tunisian insurance market and to determine various factors affecting the contact choice and accident occurrence in the studied insurance market.

For this purpose, we use a data set that includes all information that an insurer had about its policyholders for the 2009 year. Our paper differs from earlier studies in two points;

First, we use a recent data from the Tunisian insurance market, which is not investigated in the empirical literature. This data is concerning the year 2009, after the implementation of the new no-claims bonus class. Second, we use the two-stage regression analysis developed by Richaudeau (1999) where the probit model is used for coverage choice equation (first stage) and the negative binomial model for the accident frequency equation (second stage).

The remainder of this paper is organized as follows. The next section presents briefly the Tunisian automobile insurance market. Then, section 3 presents the models used and the empirical test for the presence of asymmetric information. Section 4 describes the data and reports a descriptive statistics analysis of the data set. Section 5 analyses and discusses the empirical results. The last section concludes and discusses possible future work.

2. The Tunisian Automobile Insurance Market

In Tunisia, the insurance sector plays an important role in the development of the private sector and the modernization of the securities markets[2]. For instance, the Tunisian market characterized by the presence of 22 insurance companies operating in 24 regions. We find 12 of these companies engage in multi-branch operations, 6 are single branch companies (3 specialize in life insurance, 1 in expert insurance, 1 in credit insurance and 1 in reinsurance) and 4 are off-shore companies authorized to conduct operations with nonresidents.

The insurance industry is dominated by the private sector, with 61.27% in 2010. The level of average premium has increased from 98,355 TND (in 2009 to 106,195 TND in 2010).

Looking to their products' distribution in 2008, insurance companies use about 697 intermediaries (614 insurance agents, 55 brokers and 28 agents authorized to sell life insurance) forming a network making insurance policies presents in all different regions. They employ also about 10 actuaries and 1008 experts and commissioners of damage authorized to assess and evaluate the losses and damages.

In 2010, the Tunisian market structure has not changed; the automobile insurance contributes to total business with 46.18% of the market. The life insurance branch, with 14.45%, follows it. The group health insurance, various risks, transport and fire insurance branches experienced the highest growth rates, with 13.05%, 11.22%, 6.76% and 5.65%, respectively.

[1] *Beginning drivers are those having less than 3 years' driving experience before contracting with the insurer.*

[2] *The description of the insurance market in this section is based on the notes of the F.T.U.S.A. (2010)*

3. Methodology

To investigate the presence of adverse selection in the Tunisian automobile insurance market, we employ two models: the two-stage regression analysis developed by Richaudeau (1999) and the bivaiat bivariate probit model introduced by Chiappori and Salanié (2000) and Dionne et al. (2001).

3.1 The two-stage regression analysis of Richaudeau (1999)

In this modeling, the first stage aims to estimate the coverage choice and the second stage estimates the accident occurrence.

3.1.1 Choice of a coverage equation (first stage)

In this stage, the coverage choice can be defined as a dichotomous variable:

$$C_i = \begin{cases} 1 & if \quad C_i^* = \beta X_i + \varepsilon_i \succ 0, \\ 0 & \quad otherwise, \end{cases}$$

A maximum likelihood probit regression is employed to estimate the deductible choice equation.

Then, the generalized residual is given by:

$$\hat{\varepsilon}_i = \frac{\phi(X_{1i}\hat{\beta}_1)}{\Phi(X_{1i}\hat{\beta}_1)(1-\Phi(X_{1i}\hat{\beta}_1))}\left[C_i - \Phi(X_{1i}\hat{\beta}_1)\right], \quad (3)$$

Where ϕ (.) and Φ (.) are, respectively, the density and cumulative distribution functions of the standard normal distribution, and $\hat{\beta}_1$ is an estimated coefficient vector. If the $\hat{\varepsilon}_i$ is positive (negative), then the policyholder has subscribed a third party coverage policy.

3.1.2 The accident Frequency equation (second stage)

In the second stage, we estimate the probability of accident using a count data model, and more precisely, a negative binomial model of the number of accidents. To test for asymmetric information, the residual obtained from the binary choice coverage model in equation (3) is used as input when estimating the negative binomial model[3]. The probability of having yi accidents is given by:

$$P(Y = y_i) = \frac{\Gamma\left(y_i + \frac{1}{\sigma^2}\right)\left[\sigma^2 \exp\left(X_i\hat{\beta}_2 + \hat{\varepsilon}_i\hat{\beta}_\varepsilon\right)\right]^{y_i}}{\Gamma\left(\frac{1}{\sigma^2}\right)\Gamma(y_i+1)\left[1+\sigma^2 + \exp\left(X_i\hat{\beta}_2 + \hat{\varepsilon}_i\hat{\beta}_\varepsilon\right)\right]^{y_i+\frac{1}{\sigma^2}}} \quad (4)$$

Where Γ (.) is the Gamma function, $\hat{\beta}_2$ and $\hat{\beta}_\varepsilon$ are estimated coefficients vectors. The regression coefficient $\hat{\beta}_\varepsilon$ captures the correlation between the insurance coverage and the frequency of accidents, conditional on the effects of all observables on both the coverage choice and the accident frequency. If $\hat{\beta}_\varepsilon$ is positive and statistically significant, we conclude the existence of asymmetric information. This will be our test of the efficiency in the Tunisian automobile insurance market.

The Richaudeau's test may induce an estimation bias in the second-stage equation. Using the result of Murphy and Topel (1985), we correct this bias by testing the statistical significance of $\hat{\beta}_\varepsilon$. In this last test, the asymptotic variance-covariance matrix for two-stage maximum likelihood model is defined as:

$$\sum = R_2^{-1} + R_2^{-1}[R_3^{'}R_2^{-1}R_3 - R_4^{'}R_2^{-1}R_3 - R_3^{'}R_2^{-1}R_4]R_2^{-1} \quad (5)$$

Where,

$$R_1 = E\frac{\partial L_1}{\partial \beta_1}\left(\frac{\partial L_1}{\partial \beta_1}\right)' \qquad R_2 = E\frac{\partial L_2}{\partial \beta_2}\left(\frac{\partial L_2}{\partial \beta_2}\right)'$$

$$R_3 = E\frac{\partial L_2}{\partial \beta_1}\left(\frac{\partial L_2}{\partial \beta_2}\right)' \qquad R_4 = E\frac{\partial L_1}{\partial \beta_1}\left(\frac{\partial L_2}{\partial \beta_2}\right)'.$$

[3] Based on the Akaike's Information Criterion (AIC) statistic, the negative binomial regression fits our data better than the Poisson regression (see statistics for dispersion parameters in Table 4). See also Kim et al. (2009) who show similar results for Korean data and Dionne and Vanasse (1992) for Canadian data.

3.2 Bivariate probit model

To test the conditional independence between the contract choice and the accident occurrence, we follow the empirical approach proposed by Chiappori and Salanié (2000). In fact, they use a bivariate probit model to examine the coverage-risk correlation.

In this modeling, the accident occurrence Y is presented as follows:

$$Y_i = \begin{cases} 1 & \text{if the driver incurs at least one accident during the calandar year} \\ 0 & \text{Otherwise} \end{cases}$$

Hence, the bivariate model is composed by two equations:

- *The equation of the deductible Choice*

$$C_i = \begin{cases} 1 & \textit{if} \quad C_i^* = \beta X_i + \varepsilon_i \succ 0, \\ 0 & \textit{otherwise}, \end{cases} \tag{6}$$

- *The equation of accident Occurrence*

$$Y_i = \begin{cases} 1 & \textit{if} \quad Y_i^* = \gamma X_i + \eta_i \succ 0, \\ 0 & \textit{otherwise}; \end{cases} \tag{7}$$

In this specification, β and γ are coefficients to be estimated. X_i represents the matrix of observable independent variables. The two error terms, ε and satisfy the following standard conditions:

$$E[\varepsilon_i] = E[\eta_i] = 0, \qquad Var[\varepsilon_i] = Var[\eta_i] = 1 \qquad \textit{and} \qquad Cov[\varepsilon_i, \eta_i] = \rho.$$

In this model, ρ measures the correlation between coverage and risk and allow as to conclude about the presence of residual adverse selection in the portfolio.

4. Empirical Methodology

4.1 Data

To investigate the asymmetric information, we consider a portfolio of 54040 policies that contains information for 31125 policyholders

In this empirical study, we employ 11 variables, consisting of the policyholder's car's and accident's characteristics. The description of different variables are presented in the appendix (1).

In our analysis, the coverage choice measured by third party coverage versus comprehensive coverage rather than low premium versus high premium. Thus, we define it as follows:

$$CoverageChoice = \begin{cases} 1 & \text{if the contract corresponds to a Third Party Liability} \\ 0 & \text{otherwise,} \end{cases}$$

4.2 Summary statistics

To begin with, let us give an overview of our data set.

The file contains 5301 women (17.03% of total population) and 25824 males (82.97%) of which 15742 leading vehicle for commuting and family use and 15383 driving vehicles for commercial use. The average premium is 314TND and the maximum premium that the insurer receives is 5909 TND. The number of accidents stretches from no accidents up to 8 accidents, where the mean value is 0.20 indicating a low frequency of accidents.

Most vehicles are of French brand 18306 (or 58.81% of all cars), with 9216 cars for private use and 9090 cars for commercial use, and which are respectively 50.34% and 49.66%. Moreover, about 30% of the drivers live in the Coast of Tunisia especially in the governorate of Sousse. It is very eventful on the economic, administrative and social capital. Concerning the job of the policyholders' variable, employees represent about 50% of the population and 21% are functionaries whereas 12% are artisan.

In our sample of 31125 contracts, we have information on the number of responsible accidents reported by the insured to his insurer in 2009. Our data includes 822 contracts (1176 contracts) that have at least one accident during the contract year for the beginning drivers (experienced drivers), 620 (745) contracts claimed two accidents, 114 (161) contracts made 3 accidents and 73 (123) contracts claimed more than three accidents (See table1). No single policyholder in our portfolio incurred more than eight accidents during the year.

Table 1. Distribution of number of accidents

Number of accidents	Total		Beginners drivers		Experienced drivers	
	Number	Percent	Number	Percent	Number	Percent
0	27291	87.68	13424	89.18	13867	86.28
1	1998	6.42	822	5.46	1176	7.32
2	1365	4.39	620	4.12	745	4.64
3	275	0.88	114	0.76	161	1.00
4	134	0.43	56	0.37	78	0.49
5	28	0.09	10	0.07	18	0.11
6	19	0.06	5	0.03	14	0.09
7	6	0.02	0	0.00	6	0.04
8	9	0.03	2	0.01	7	0.04
Total	31125	100	15053	100	16072	100

5. Results

5.1 The choice of coverage

Table 2 presents the estimation results of the probit modeling that is the first step regression analysis of Richaudeau (1999). The left-hand side of this table suggests the estimation for the beginning drivers whereas the right-hand side presents the estimation for the experienced drivers.

For the beginning drivers, commuting and family use cars are more likely to be insured by comprehensive insurance than commercial use cars. In addition, we note that females have a greater probability of choosing a comprehensive coverage than males, indicating perhaps that their risk-aversion is greater. The probability of buying third party insurance is higher for the policyholders who have age between 56 and 75 years.

For the experienced drivers, the probability of choosing third party coverage is more important for female and for the policyholders who are older than 31 years. Primary use cars are more likely to be insured by comprehensive coverage than commercial use cars.

One interesting point to note is that the jobs of the policyholders also have a role in the choice of coverage. The probability of choosing third party coverage decreases with the senior executive, the artisan, the employee and the unemployed.

Table 2. Probit model results

Independent variables	Beginning drivers		Experienced drivers	
	Coefficient	p-Value	Coefficient	p-Value
Female	0.215	0.000	-0.213	0.000
Primary use	-0.040	0.069	0.065	0.008
Age 2	-0.122	0.565	-0.294	0.212
Age 3	-0.237	0.258	-0.771	0.001
Age 4	-0.520	0.014	-0.716	0.002
Age 5	-0.203	0.361	-0.486	0.039
Job 1	0.065	0.152	-0.162	0.002
Job 2	0.093	0.227	-0.069	0.425
Job 3	0.145	0.101	-0.260	0.007
Job 4	0.015	0.766	-0.084	0.145
Job 5	0.002	0.952	-0.113	0.020
Job 6	0.039	0.526	-0.263	0.000
Origin 1	-0.024	0.861	0.009	0.949
Origin 2	-0.012	0.931	0.030	0.832
Origin 3	-0.030	0.832	-0.006	0.966
Origin 4	-0.022	0.878	0.013	0.925
Origin 5	-0.148	0.502	0.308	0.297
Origin 6	0.011	0.966	0.516	0.108
Zone 1	0.030	0.418	-0.023	0.557
Zone 2	0.044	0.314	0.024	0.623

Zone 3	0.008	0.852	0.002	0.96
Zone 4	-0.032	0.380	-0.035	0.397
Zone 5	0.006	0.905	0.062	0.318
Constant	0.851	0.001	1.922	0.000
Number of obs	15053		16072	
Log likelihood	-8618.224		-6764.429	

5.2 The probability of having an accident

According to the estimated parameters in the second equation related to the occurrence of accidents (the second part of table 4), the probability of making at least one accident at the insurer increases according to female and to the agent lives in big Tunis. Furthermore, this probability decreases with job of the drivers. Particularly, the accident occurrence is negatively related to the unemployed person, the employee and the retired person.

Table 3 shows that many exogenous variables have a significant effect on accidents frequency for both beginning and experience drivers. A special note is the need for some exogenous variables. First, we examine the frequency of accidents for beginning drivers. The positive sign for the coefficients of gender and the residence area that suggests that female and policyholders live in big Tunis have a higher chance to incur accidents. In addition, we observe a positive effect of vehicle's use on the probability of accident. The policyholder specifies the primary use of the insured car as for commercial use that has a higher chance of accidents. Second, we move to examine the accident frequencies for experienced drivers. We observe that the accident frequency is positively related to the job and the residence area of policyholders. This probability decreases according to the age of drivers.

Table 3. Negative Binomial model Results

Independent variables	Beginning drivers			Experienced drivers		
	Coefficient	Mtopel Std.Err.	p-Value	Coefficient	Mtopel Std. Err.	p-Value
	Accident frequencies equation					
Female	1.377	0.943	0.144	0.253	0.324	0.434
Primary use	-0.347	0.203	0.088	0.062	0.099	0.529
Age 2	-0.729	0.813	0.370	-1.319	0.305	0.000
Age 3	-1.369	1.189	0.250	-0.902	0.744	0.225
Age 4	-3.729	2.486	0.134	-1.776	0.674	0.008
Age 5	0.192	1.076	0.858	0.176	0.421	0.676
Job 1	0.759	0.341	0.026	0.818	0.243	0.001
Job 2	0.966	0.484	0.046	0.593	0.186	0.001
Job 3	1.201	0.701	0.087	0.791	0.409	0.054
Job 4	0.561	0.210	0.789	0.824	0.160	0.000
Job 5	-0.420	0.157	0.008	0.270	0.181	0.134
Job 6	-0.264	0.323	0.413	0.418	0.387	0.280
Origin 1	-0.123	0.562	0.827	0.514	0.326	0.115
Origin 2	-0.039	0.567	0.945	0.501	0.336	0.136
Origin 3	0.451	0.570	0.429	0.993	0.329	0.003
Origin 4	-0.682	0.589	0.247	0.772	0.332	0.020
Origin 5	0.351	0.973	0.718	0.747	0.607	0.218
Origin 6	0.569	0.844	0.501	1.202	0.736	0.102
Zone 1	1.013	0.224	0.000	0.946	0.091	0.000
Zone 2	0.292	0.283	0.303	0.007	0.113	0.949
Zone 3	0.487	0.220	0.027	0.287	0.107	0.008
Zone 4	0.095	0.228	0.677	0.269	0.101	0.008
Zone 5	0.082	0.289	0.775	0.337	0.147	0.022
Constant	11.330	11.616	0.329	-3.149	5.812	0.588
RESIDU	-16.742	14.483	0.248	1.065	5.828	0.855
Lnalpha cons	1.519	0.062	0.000	1.229	0.472	0.000
Number of obs	15053			16072		
Log likelihood	-6680.201			-8486.713		

5.3 Testing for asymmetric information

According to table 4, the first estimation results show that the correlation between the error terms of the two binary equations is positive. Thus, the estimated value of ρ is 0.126 for beginning drivers. This association is strong and significant with a p-value of 0.000. Hence, we find that policyholders who purchase better coverage tend to be more prone to make claims. We therefore, find an evidence of asymmetric information, either with reference to adverse selection or with reference to moral hazard, especially when the observed variables are taking into account

The second estimation results show that we again find a positive correlation between coverage-risk for experienced drivers. The estimates value of ρ is 0.001. The p-value of ρ, however, suggests that this association is not statistically significant (ρ= 0.963). Therefore, we find no evidence of adverse selection or moral hazard. This result is consistent with that of Saito (2006).

Table 3 provides, also, the estimate of the risk-coverage correlation for beginning drivers and experienced drivers. The negative binomial model gives us a similar result as the probit bivariate model. For the beginning drivers, the coefficient of the coverage choice residual variable is positive and statistically significant, this finding clearly demonstrates that asymmetric information is present in the Tunisian automobile insurance market. However, this coefficient becomes negative and statistically insignificant for the experienced drivers.

To sum up, all these results show no evidence of asymmetric information for experienced drivers and the existence of evidence for beginning drivers.

Table 4. Bivariate Probit Estimation Results

Independent Variables	Beginning drivers		Experienced drivers	
	Coefficient	p-Value	Coefficient	p-Value
	Choice of contractequation			
Female	0.216	0.000	-0.213	0.000
Commercial use	0.040	0.068	0.065	0.008
Age 1	0.207	0.353	-	-
Age 2	0.082	0.308	-0.294	0.212
Age 3	-0.033	0.666	-0.771	0.001
Age 4	-0.316	0.000	-0.716	0.002
Age 5	-	-	-0.486	0.039
Job 1	0.065	0.154	-0.162	0.002
Job 2	0.092	0.230	-0.069	0.425
Job 3	0.145	0.101	-0.260	0.007
Job 4	0.014	0.782	-0.084	0.145
Job 5	0.001	0.981	-0.113	0.020
Job 6	0.038	0.538	-0.263	0.000
Origin 1	-0.024	0.864	0.009	0.949
Origin 2	-0.011	0.938	0.030	0.833
Origin 3	-0.029	0.836	-0.006	0.966
Origin 4	-0.022	0.879	0.013	0.925
Origin 5	-0.149	0.501	0.308	0.297
Origin 6	0.001	0.998	0.516	0.108
Zone 1	0.030	0.419	-0.023	0.575
Zone 2	0.043	0.322	0.024	0.623
Zone 3	0.008	0.854	0.002	0.961
Zone 4	-0.032	0.378	-0.035	0.397
Zone 5	0.007	0.905	0.062	0.318
Constant	0.606	0.000	1.922	0.000
	Accident occurrenceequation			
Female	0.204	0.000	0.078	0.047
Commercial use	0.107	0.000	-0.032	0.227
Age 1	-0.759	0.001	-	-
Age 2	-1.032	0.000	-1.089	0.000
Age 3	-1.043	0.000	-0.946	0.000
Age 4	-1.426	0.000	-1.404	0.000
Age 5	-	-	-0.056	0.683

Job 1	0.228	0.000	0.415	0.000
Job 2	0.269	0.000	0.347	0.000
Job 3	0.221	0.002	0.489	0.000
Job 4	-0.084	0.026	0.428	0.000
Job 5	-0.249	0.198	0.085	0.122
Job 6	-0.294	0.000	0.234	0.002
Origin 1	-0.020	0.001	0.257	0.144
Origin 2	-0.007	0.917	0.231	0.198
Origin 3	0.339	0.973	0.537	0.002
Origin 4	-0.323	0.081	0.433	0.016
Origin 5	0.783	0.108	0.559	0.039
Origin 6	0.308	0.002	0.668	0.016
Zone 1	0.462	0.312	0.540	0.000
Zone 2	0.040	0.000	-0.017	0.774
Zone 3	0.172	0.532	0.109	0.061
Zone 4	0.121	0.007	0.121	0.013
Zone 5	-0.054	0.024	0.219	0.001
Constant	-0.521	0.015	-0.989	0.000
Athrho	0.127	0.000	0.001	0.963
Rho	0.126	0.000	0.000	0.963
Number of obs	15053		16072	
Log likelihood	-13226.16		-12483.116	
Wald chi2(46)	1168.53		1482.76	
Prob>chi2	0.000		0.000	
Like lihood-ratio test of rho=0				
	$\chi^2(1) = 39.700(P > \chi^2 = 0.000)$		$\chi^2(1) = 0.002(P > \chi^2 = 0.963)$	

6. Conclusion

In this paper, we examined coverage-risk relationship in the Tunisian automobile insurance market using a cross sectional data from a major insurer. We used the bivariate probit model and the negative binomial model to help describing any evidence of private information. When analyzing the entire portfolio of policyholders from the insurer, all our procedures give the same results. For the beginning drivers, we found significant positive relationship between policyholder's coverage choice and the occurrence of accidents, indicating the presence of adverse selection. This relationship provides evidence that some unobserved and unobservable factors affects the insurance contract choice. Therefore, for the experienced drivers, asymmetric information seems to be at most a negligible phenomenon in the insurance automobile market.

One major criticism of the conditional correlation approach with cross sectional data is that it does not allow separation adverse selection from moral hazard. Our analysis is limited to a combined result of the two asymmetric information problems. However, as suggested by Abbring et al. (2003), dynamic data could help distinguish between the adverse selection and moral hazard. Finally, we leave this as one of the future research directions to pursue.

Appendix 1.

Characteristics of the Driver

Gender: 1 if the insured is male (82.97%); 0 if the insured is female (17.03%).

Age: the age of the policyholder (mean =45.49, std. dev. =13.29, min= 18, max= 98).

Age's categories : we use five dummy variables

Age 1 = 1 if the policyholder have age between 18 and 21 years (0.47%); 0 otherwise.

Age 2 = 1 if the policyholder have age between 22 and 30 years (11.60%); 0 otherwise.

Age 3 = 1 if the policyholder have age between 31 and 55 years (66.86%); 0 otherwise.

Age 4 = 1 if the policyholder have age between 56 and 75 years (16.65%); 0 otherwise.

Age 5 = 1 if the policyholder is older than 75 years (4.42%); 0 otherwise.

Job: For the job of the policyholder, many possible classifications can be used to distinguish between professions for policyholders. In this study, we propose seven dichotomous variables defined as follows:

Job 1 = 1 if the policyholder is an official (21.21%); 0 otherwise.
Job 2 = 1 if the policyholder is a senior executive (2.78%); 0 otherwise.
Job 3 = 1 if the policyholder is an unemployed person (8.01%); 0 otherwise.
Job 4 = 1 if the policyholder is a craft (11.96%); 0 otherwise.
Job 5 = 1 if the policyholder is an employee (48.61%); 0 otherwise.
Job 6 = 1 if the policyholder is a retired person (5.49%); 0 otherwise.
Job 7 = 1 if the policyholder is a middle manager (1.94%); 0 otherwise.

- Residence area variables: we define 6 dichotomous variables to capture the effect of the immediate environment of the individual driver.
Zone 1 = 1 if the policyholder lives in big Tunis (28.35%); 0 otherwise.
Zone 2 = 1 if the policyholder lives in North (12.45%); 0 otherwise.
Zone 3 = 1 if the policyholder lives in Northwest (10.83%); 0 otherwise.
Zone 4 = 1 if the policyholder lives in Coast (29.90%); 0 otherwise.
Zone 5 = 1 if the policyholder lives in Center (5.71%); 0 otherwise.
Zone 6 = 1 if the policyholder lives in South (12.76%); 0 otherwise.

Characteristics of the car
- Use of the car = 1 if the policyholder specifies the primary use of the insured car as for commuting and family use (50.58%); 0 otherwise (i.e., commercial use) (49.42%).

Car brand: We define seven categories that capture the car's country-of-origin effect.
Origin 1 = 1 if the car made in France (58.81%); 0 otherwise.
Origin 2 = 1 if the car made in Italy (12.62%); 0 otherwise.
Origin 3 = 1 if the car made in German (17.31%); 0 otherwise.
Origin 4 = 1 if the car made in China or in Korea or in Japan (9.92%); 0 otherwise.
Origin 5 = 1 if the car made in the USA (0.34%); 0 otherwise.
Origin 6 = 1 if the car made in England (0.28%); 0 otherwise.
Origin 7 = 1 if the car made in other countries (0.70%); 0 otherwise.
Past involvement in accidents
Accidents' number: the number of accident (mean = 0.207, std. dev. = 0.646, min = 0, max=8).
Bonus-malus classes: it is a discrete variable. There are 8 classes from 1 to 8 (mean =5.845, std. dev. =2.275, min=0, max=8).
Indemnity: it is a continue variable (mean = 163, std. dev. = 1511.232, min = 0, max = 105111).

References

Abbring J., Heckman J., Chiappori P.A., and Pinquet J. (2003): "Adverse selection and moral hazard in insurance: Can dynamic data help to distinguish?" Journal of the European Economic Association, 1(2-3), pp 512-521.
Chiappori P.A. and Salanié B. (2000): "Testing for Asymmetric Information in Insurance Markets", Journal of Political Economy, 108, pp 56-78.
Cohen A. (2005): "Asymmetric information and learning: evidence from the automobile insurance market", The Review of Economics and Statistics, 87, pp 197-207.
Cohen A. and Siegelman P. (2010): "Testing for Adverse Selection in Insurance Markets", Journal of Risk and Insurance, 77(1), pp 39-84.
Cutler D.M. and Reber S.J. (1998): "Paying For Health Insurance: The Trade-Off between Competition and Adverse Selection", The Quarterly Journal of Economics, 113, pp 433–466.
Cutler D.M. and Zeckhauser R.J. (1998): "Adverse Selection in Health Insurance", Forum for Health Economics & Policy, Vol. 1: (Frontiers in Health Policy Research), Article 2.
Dahlby B. (1983): "Adverse Selection and Statistical Discrimination: An Analysis of Canadian Automobile Insurance", Journal of Public Economics, 20, pp 121-130.
Dahlby B. (1992): "Testing for Asymmetric Information in Canadian Automobile Insurance", Contributions to Insurance Economics (Boston: Kluwer Academic Publishers), pp 423-443.
Dionne G. and Ghali O. (2005): "The (1992) Bonus-Malus System in Tunisia: An Empirical Evaluation", Journal of Risk and Insurance, The American Risk and Insurance Association, 72(4), pp 609-633.
Dionne G., Gouriéroux C. and Vanasse C. (2001): "Testing for Evidence of Adverse Selection in the Automobile Insurance Market: A Comment", Journal of Political Economy, 109, pp 444-453.
Dionne G. and Vanasse C. (1992): "Automobile Insurance Ratemaking in the Presence of Asymmetrical Information", Journal of Applied Econometrics, 7, pp 149-165.
Dionne G., St-Amour P. and Vencatachellum, D. (2009): "Asymmetric information and adverse selection in Mauritian slave auctions", Review of Economic Studies, 760 (4), pp 1269-1295.

Dionne G., Michaud P.C., and Pinquet J. (2013): "A review of recent theoretical and empirical analyses of asymmetric information in road safety and automobile insurance", Research in Transportation Economics, 43, pp 85-97.

Englund M. (2010): "Asymmetric Information: Evidence from the Automobile Insurance Market", Submitted to the Journal of Risk and Insurance.

Fang H., Keane M.P. and Silverman D. (2008): "Sources of advantageous selection: Evidence from the medigap insurance market", Journal of Political Economy, 1160 (2), pp 303-350.

Fédération Tunisienne des Sociétés d'Assurance (2010), Rapport annuel, December, Tunis.

Février P., Linnemer L., and Visser M. (2012): "Testing for asymmetric information in the viager market", Journal of Public Economics, 96, pp 104-123.

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

Finkelstein A. and Poterba J. (2002): "Selection Effects in the United Kingdom Individual Annuities Market," The Economic Journal, 112, pp 28–50.

Grun-Réhomme M. and Ben Lagha N. (2007) : "Choix de contrat et sinistralité chez les jeunes conducteurs", Assurances et Gestion des Risques, 74(4), pp 505-532.

Kim H., Kim D., Im S. and Hardin J.W. (2009): "Evidence of asymmetric information in the automobile insurance market: Dichotomous versus multinomial measurement of insurance coverage", The Journal of Risk and Insurance, 76(2), pp 343-366.

Lemaire, J. (1995): "Bonus-Malus Systems in Automobile Insurance", Boston: Kluwer Academic Publishers, pp283.

Li C.S., Liu C.C., and Peng S.C. (2013): "Bundled automobile insurance coverage and accidents", Accident Analysis and Prevention, 50, pp 64-72.

Makki, S.S. and Somwaru A. (2001): "Evidence of Adverse Selection in Crop Insurance Markets", Journal of Risk and Insurance, 68(4), pp 685–708.

Puelz R. and Snow A. (1994): "Evidence on Adverse Selection Equilibrium Signaling and Cross-subsidization in the Insurance Market", Journal of Political Economy, 102, pp 236-257.

Richaudeau D. (1999): "Automobile Insurance Contracts and Risk of Accident: An Empirical Test Using French Individual Data", Geneva Papers on Risk and Insurance Theory ski, 24, pp 97-114.

Rothschild M. and Stiglitz J. (1976): "Equilibrium in Competitive Insurance Markets: an Essay on the Economics of Imperfect Information", Quarterly Journal of Economics, 16, pp 629-649.

Saito K. (2006): "Testing for Asymmetric Information in the Automobile Insurance Market under Rate Regulation", Journal of Risk and Insurance, 73(2), pp 335-356.

Sil J. (2005): "Endogenous Adverse Selection: Evidence from U.S. Crop Insurance," Mimeo, University of California, Berkeley, Department of Agricultural and Resource Economics, available at http://are.berkeley.edu/~jsil/jobmarket.html.

Shi P. and Valdez E.A. (2011): "A copula approach to test asymmetric information with applications to predictive modeling", Insurance: Mathematics and Economics, 49, pp 226-239.

Stiglitz, J. (1977): "Monopoly, non-linear pricing, and imperfect information: the insurance market", Review of Economic Studies, 44 (3), 407-430.

Wilson, C. (1977): "A model of insurance markets with incomplete information", Journal of Economic Theory, 16 (2), 167-207.