



AN APPLICATION OF CLAIM FREQUENCY DATA USING ZERO INFLATED AND HURDLE MODELS IN GENERAL INSURANCE

DOI: 10.17261/Pressacademia.2015414539

Latife Sinem Sarul¹, Serap Sahin²

¹Istanbul University, School of Business Quantitative Methods Department, lasinem@istanbul.edu.tr

² Kirikkale University, School Economics and Administrative Sciences, shnsrp@gmail.com

Keywords

Insurance, Claim
Frequency Data,
Zero Inflated
Models, Hurdle
Models

ABSTRACT

Modelling is a key issue to get a fair pricing in insurance. Poisson distribution is the basic model for count data when the assumptions of Poisson process are assured. Since the insurers tend not to state the small claims to get the deductibles and no claim discounts, insurance data has more zeros than expected which makes the contradictions of the Poisson process assumptions that is the equality mean and variance value. However, not to take account excess zeros makes the knowledge deficiency to get the better pricing for the portfolio. In this paper we will compare Poisson Models and Zero Inflated Models which account for this fact for claim frequency data. Also we will review the models in use for count data and also compare Hurdle Models as an alternative to Zero Inflated Models. Our results represent that Hurdle Models are better fit than the other models we compare. We used Akaike's Information Criteria(AIC) as model selection measures.

JEL Classification
C46, C10, D30

1. INTRODUCTION

Modelling is a key issue in insurance to get a fair pricing which is the major task for the actuaries. In this concept, the distribution of the data set which is investigated is very important. Poisson distribution is the basic model to be applied to count data. However, the equivalence of mean and variance assumption of the Poisson process is not satisfied for insurance data because the insurers tend not to state the small claims to get the deductibles and no claim discounts which cause insurance data set generally has more zeros than expected. In that reason overdispersion which is the situation of that the variance is greater than the mean is seem to be main problem for the insurance dataset. Negative Binomial Model is developed as an alternative to handle overdispersed count data. However, the Negative Binomial Model can also be violated by overdispersion when the variance is greater than the calculated value of $\mu + \alpha\mu^2$. Overdispersion can be occurred in two ways as apparent and real overdispersion. Missing variables, outliers, requiring interaction term or misspecified link function can cause apparent overdispersion

while the violation of distributional assumptions can cause real overdispersion. Zero Inflated Models developed to handle inflated zero values for the dependent variable which cause the violation of distribution assumptions (Hilbe,2012).

2. LITERATURE REVIEW

Zero Inflated Poisson Models are developed by Lambert (1992) to handle zero-inflated count data. Zero Inflated Models combine two sources of zero outcomes which are called "true zeros" and "excess zeros". Greene(1994) has investigated Zero Inflated Models as modifications of the Poisson and the Negative Binomial models. He also presents the test procedure to separate the zero inflation and overdispersion. Fahrmeir and Echavarría (2006) developed structured additive regression models for overdispersed and zero inflated data. They applied the models to the patent data and also motor insurance data to investigate good indicators for patent. Boucher, Denuit and Guillen (2009) presented different risk classification models for panel count data based on the Zero Inflated Poisson distribution. They suggested a new approximation taking account the behavior of insureds. Zhao&Zhou (2012) discussed longitudinal models of claim counts with excess zeros. They modelled claim counts by using copula function.

Mullahy (1986) has first discussed in the econometric literature hurdle count data models which is also called two part models by Heilbron (Heilbron, 1994). Gurmu (1998) introduced a Generalized Hurdle Model for the handling of overdispersion and also underdispersion which is the situation of that the variance is less than the mean. Ridout et al.(1998) reviewed some zero inflated models and hurdle models and gave an example on biological count data. Saffari, Adnan and Greene (2012) suggested using a Hurdle Negative Binomial Regression Model to overcome the problem of overdispersion. They introduced a censored Hurdle Negative Binomial Model on count data with many zeros. Greene (2007) has compared Zero Inflated and Hurdle Models. In this work it is also described several extensions of the models and is presented an application to health care demand data for comparison the models. Hurlimann (1990) presented the usage of parametric models on the claim frequency distribution with extra zeros. Yip&Yao (2006) present application of Zero Inflated Models to insurance claim frequency data. They reviewed the development of Zero Inflated Models and take attention that there is very few application on claim frequency data in the literature. Flynn (2009) compared traditional Poisson and Negative Binomial models to the Zero Inflated Models. He applied Zero Inflated Models and some data mining techniques to claim frequency data and also attached efficient preprocess procedure for the categorical variables. Morata (2009) examined a priori ratemaking procedure for the portfolios include different types of claims by using bivariate Poisson Models with Zero Inflated Models. Bermudez and Karlis (2011) extended this work based on Bayesian inference by using multivariate Poisson Regression Models with their zero inflated versions. Mouatassim and Ezzahid(2012) compare poisson model to the zero inflated model and applied to health insurance data set. Mouatassim, Ezzahid and Belasri(2012) analyzed operational risk to the zero inflated data and assess the impact of ZIP distribution on the operational capital charge. They concluded that the zero inflated Poisson distribution is better fit then Poisson distribution for modeling operational risk frequency.

In this paper, we are motivated from Yip and Yau (2006) and try to extend their results with different data set. We used a company specific data with four independent variables. In the next section, we give a brief literature and then we continued with quick explanation about traditional Poisson and Negative Binomial Models beside Zero Inflated and Hurdle Models. We give descriptive statistics and counts for our dependent variable which is claim frequency to monitor our data set. Finally we present and summarized our results in the last section.

3. DATA AND METHODOLOGY

Data

The data set that we considered in this paper contains claim frequencies for automobile portfolio of a Turkish insurance company occurred between 2012 and 2014. The data contains information from 10.814 policyholders and it is a company-specific data. There are four independent variables which are insureds gender, marital status, age and vehicles age. Below Figure1 presents histogram of claim frequencies of policyholders on the left hand side. The histogram of Claim Frequencies represents that the dataset contains very large proportion of zeros which refers to overdispersion. Claim counts are also seen on the right hand side in Table1.

Figure 1: Histogram of Claim Frequency

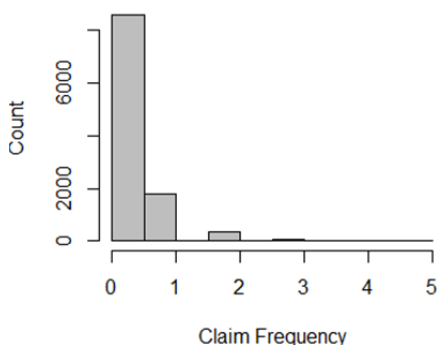


Table 1: Counts of Claim Frequencies

Claim Frequency	Count
0	8544
1	1796
2	370
3	81
4	22
5	1

According to the Table1 given above, the data set contains 8.544 zero counts which is 79% of the data set. Also, 3 and greater than 3 claim frequencies form approximately 1% of the data set. Below it is given the distribution of counts and ratios according to the explanatory variables by codes in Table2. According to the Table 2, looking at the insureds characteristics, 72% of the data set are male and 28% of the data set are female. In addition, 78% are married, 15% are single, 5% are divorced and 2% of the data set are widow. Furthermore, insureds aged between 36 and 50 years filed about 38% of total claims and 31% of the insureds are younger than 35. And, 3% of the insureds are older than 71 years old. Looking at the vehicles properties, it is seem that 80% of the vehicles are between 0 and 5 years old, 19% of the vehicles are between 6 and 10 years old, 1% of the vehicles are between 11 and 15 years old.

Table 2: Summary of the dataset

Variables	Code	Groups	Claim Frequency	Ratio
Gender	1	Female	3047	28
	2	Male	7767	72
Marital Status	1	Single	1617	15
	2	Married	8457	78
	3	Divorced	503	5
	4	Widow	237	2
Vehicle Age	1	0-5	8637	80
	2	6-10	2011	19
	3	11-15	155	1
	4	16-20	11	0
Insureds Age	1	0-35	3397	31
	2	36-50	4108	38
	3	51-70	2940	27
	4	71-100	369	3

Methodology

In this section, we first lay out the Probability Mass Function (PMF) for the Poisson and Negative Binomial model. Then we continued with the same for Zero Inflated Poisson, Zero Inflated Negative Binomial and the Hurdle Model. Poisson and Negative Binomial distribution is traditionally used the claim count distribution in general insurance.

Poisson distribution

The Poisson probability distribution function is formulated as

$$P(Y = y|\lambda) = \frac{e^{-\lambda}\lambda^y}{y!} \quad y = 0, 1, 2, \dots$$

λ is the expected value and the variance of Poisson distribution.

Likelihood function for the Poisson Model:

$$L(\beta|y, X) = \prod_{i=1}^N P(Y = y_i|\lambda_i) = \prod_{i=1}^N \frac{e^{-\lambda_i}\lambda_i^{y_i}}{y_i!}$$

where $\lambda_i = E(y_i|x_i) = e^{x_i\beta}$.

Poisson Regression Model provide a standard framework for the analysis of count data with unignorable assumptions. One of the assumptions of the Poisson distribution is the mean and the variance equivalence. In practice, however, count data often exhibit larger

variance than predicted by the mean (overdispersion). Therefore count data will be more convenient to use the Negative Binomial distribution is not provided on the assumption of the Poisson distribution (Hardin&Hilbe, 2011) (Loeys et al., 2012).

Negative Binomial Distribution

$$P(Y = y|\lambda, \alpha) = \frac{\Gamma(y + \alpha^{-1})}{y! \Gamma(\alpha^{-1})} \left[\frac{\alpha^{-1}}{\alpha^{-1} + \lambda} \right]^{\alpha^{-1}} \left[\frac{\lambda}{\alpha^{-1} + \lambda} \right]^y$$

λ is the expected value and the variance of the distribution. α is the over dispersion parameter, when $\alpha = 0$ the Negative Binomial distribution is the same as Poisson distribution. Likelihood function for the Negative Binomial Model:

$$L(\beta|y, X) = \prod_{i=1}^N P(y_i|x_i) = \prod_{i=1}^N \frac{\Gamma(y_i + \alpha^{-1})}{y_i! \Gamma(\alpha^{-1})} \left[\frac{\alpha^{-1}}{\alpha^{-1} + \mu_i} \right]^{\alpha^{-1}} \left[\frac{\mu_i}{\alpha^{-1} + \mu_i} \right]^{y_i}$$

where $\mu_i = E(y_i|x_i) = e^{x_i\beta}$ and parameter values of count models are estimated using maximum likelihood. One of the most common specifications for Negative Binomial Model is to allow the ratio of the variance to the expected value of Y to vary according to the following specification:

$$\frac{Var(Y_i)}{E(Y_i)} = 1 + \alpha E(Y_i)$$

where $\alpha = \ln(\sigma)$ and σ is the "over dispersion parameter". Under this specification, when $\alpha = 0$ (i.e. $\sigma = 1$) the model reduces to the Poisson. Because the Negative Binomial and Poisson models are nested in this way, t-tests for $\sigma = 1$, or a Likelihood-Ratio or Wald test, can be used to test for the presence of significant amounts of over dispersion (Zorn, 1996).

Zero-Inflated Models

The other problem with Poisson Regression Models having far more zeros than expected by the distributional assumptions of the Poisson and Negative Binomial models result in incorrect parameter estimates. (Hardin& Hilbe,2012) Using Zero Inflated Poisson (Lambert, 1992) or Zero Inflated Negative Binomial models are proposed as a solution for this problem (Loeys et al., 2012).

In general, count responses having two kinds of zeros "true zeros" and "excess zeros", at that point Zero Inflated Models attempt to account for excess zeros. Zero Inflated Models estimate two equations, one for the count model and one for the excess zeros.

$$P(Y = y|\lambda, \pi) = \begin{cases} \pi + (1 - \pi)e^{-\lambda} & , y = 0 \\ (1 - \pi) \frac{e^{-\lambda}\lambda^y}{y!} & , y = 1, 2, \dots \end{cases} \quad (1)$$

λ is the expected value of Poisson distribution and π is the over dispersion parameter.

Hurdle Models

Hurdle Models were developed by Mullahy(1986) to deal with count responses having more zeros than allowed by the distribution assumptions of Poisson and Negative Binomial regression, like Zero Inflated models. (Hardin& Hilbe,2012) The differences between the Hurdle Models and the Zero Inflated Models are that zero and non-zero counts are separately modelling in the Hurdle Models(Loeys et al., 2012) and also Hurdle model assumes that all zero counts are true zeros(Potts&Elith, 2006).

The Hurdle Model of count data can be expressed as follows for the Poisson and Negative Binomial distribution.

The Poisson Hurdle Model Specification

We consider a Hurdle Poisson Regression Model in which the response variable Y has the distribution

$$P(Y = y|\lambda, \pi_0) = \begin{cases} \pi_0 & y = 0 \\ \frac{(1 - \pi_0) e^{-\lambda} \lambda^y}{(1 - e^{-\lambda})y!} & y > 0 \end{cases}$$

Zero inflated Poisson distribution is parameterized as given by Equation (1) with $\pi_0 = \pi + (1 - \pi)e^{-\lambda}$.

The Negative Binomial Hurdle Model Specification

We consider a Hurdle Negative Binomial Regression Model in which the response variable Y has the distribution

$$P(Y = y|\lambda, \alpha, \pi_0) = \begin{cases} \pi_0 & y = 0 \\ (1 - \pi_0) \frac{g}{1 - (1 + \alpha\lambda)^{-\alpha^{-1}}} & y > 0 \end{cases}$$

where $g = g(y; \lambda, \alpha) = \frac{\Gamma(y+\alpha^{-1})}{\Gamma(y+1)\Gamma(\alpha^{-1})} (1 + \alpha\lambda)^{-\alpha^{-1}-y} \alpha^y \lambda^y$.

Maximum Likelihood Estimation (MLE) method is used to estimate parameters in the Zero Inflated Models.

This study includes Poisson, Negative Binomial, ZIP, ZINB, Hurdle and Hurdle NB to accommodate the excess zeros for insurance claim count data. In this paper, Akaike's information criteria (AIC) and log-likelihood values are used for model selection measures. It is also used dispersion parameters to test for overdispersion. The generalized Pearson χ^2 statistic which is the standard measure of goodness of fit is used to evaluate the sufficiency of the analyzing methods. It has the calculated value as follows;

$$\chi^2 = \sum_{i=1}^n \frac{(Y_i - \mu_i)^2}{Var(Y_i)}$$

This statistic follows a χ^2 distribution with $n - k$ degrees of freedom, where n is the total number of observations and k is the number of parameters (Cameron&Trivedi, 1998). Akaike's information criteria (AIC) and log-likelihood are basic methods of assessing the performance of the models and model selection. The AIC is defined as follows:

$$AIC = -2 \log \text{likelihood} + 2k$$

where k = number of parameters. In general, the smaller AIC value refers to the better model.

4. RESULTS

In this section we follow analysis of our data set. We applied zero inflated and hurdle models to the insurance data set and compare to the poisson and negative binomial models. Parameter estimates of models is given in Appendix 1. Below it is given Akaike information criteria and log-likelihood values to select best model of the data set in Table 3.

Table 3: Information criteria for models

Models	AIC	Log-likelihood
Poisson	14277	-7127,4
Negative Binomial	14040	-7008,2
Zero Inflated Poisson	14068	-7011,9
Zero Inflated Negative Binomial	14052	-7003,2
Hurdle	14067	-7011,7
Hurdle Negative Binomial	14052	-7003,2

According to the Table 3, Poisson Model is not best performing model for the data set with the biggest AIC value. Since the Negative Binomial model has the smallest AIC value, one can say that Negative Binomial Model is the best model for the data set. However, dispersion parameter for the Negative Binomial Model is 1,0405 and dispersion parameter for the Poisson Model is 1,201623 which represent that dependent variable claim frequency is overdispersed. On the other hand, we used Vuong test if the zero inflated model is above the poisson model and zero inflated negative binomial model is above the negative binomial model. The Vuong test (Vuong, 1989) is a test to compare Zero inflated methods to the non-nested models for counts data. For the poisson part, computed statistic of Vuong's test is $V = -6,722759$ (p value = $8,91579e-12$) which indicates that Zero Inflated Poisson regression fits better than standard Poisson regression and for the Negative Binomial part, computed statistic of Vuong's test is $V = -1,569072$ (p value = $0,0583156$) which indicates that Zero Inflated Negative Binomial Model fits better than standard Negative Binomial Model. We can also state that zero inflated poisson model is above the standard poisson model and zero inflated negative binomial model is above the negative binomial model. Because Zero Inflated Negative Binomial Model and Hurdle Negative Binomial Model have the same AIC and log-likelihood value, we can say that

these models perform best for our data set. Parameter estimates of all models are represented below in Appendix1. Using parameter estimates in Appendix1, it is possible to interpret the variables. Male drivers were found to be involved in crashes $11.35\%(1 - e^{0,1076})$ according to the zero inflated poisson model, $11.9\%(1 - e^{0,1131})$ according to the hurdle negative binomial model, 11.04% according to the zero inflated negative binomial model less compared to female drivers. It is possible to excess the results for all covariates. Here, we would like to take attention that these models has approximately close AIC values and log-likelihood results. Looking at the covariates these models gives the approximately close results. It is more reasonable to say for the decision maker, which model is the best for the data depends on the data structure. If all of the zeros are excess zeros then hurdle model is more appealing. On the other hand, if zero counts consist of both true and excess zeros that makes zero inflated models are more appealing(Xie et.al.,2013).

5. CONCLUSION

Modelling is fundamental to get a fair pricing in insurance. In that reason, analyzing the distribution of the data set is crucial for the actuaries. Especially insurance data set has more zeros than expected refers analysts to work with zero inflated models. In this paper, poisson and negative binomial models which are traditional methods to analyze count data and zero inflated poisson, zero inflated negative binomial, hurdle model with hurdle negative binomial model are applied to automobile insurance data. There are four independent variables which are insureds gender, marital status, age and vehicles age. Plot analysis presents that the dataset contains very large proportion of zeros which refers to overdispersion. Dispersion parameter is used to see if there is oversdispersion and Vuong test is used to compare non-nested models. AIC and log-likelihood values is used to compare models. We concluded that zero inflated poisson model is superior to the standard poisson model and zero inflated negative binomial model is superior to the negative binomial model. Since zero inflated negative binomial model and hurdle negative binomial model have the same AIC and log-likelihood value, we can say that they perform best for our data set. However, it is more reasonable to say for the decision maker, which model is the best for the data is depend on the data structure. Male drivers were found to be involved in crashes 11.35% according to the zero inflated poisson model, 11.9% according to the hurdle negative binomial model, 11.04% according to the zero inflated negative binomial model less compared to female drivers. It is possible to excess the results for all covariates. As a future research we intend to extend the analysis with longitudinal data to compare several zero models for insurance data.

REFERENCES

- Bermudez L. and Karlis D. , 2011, "Bayesian multivariate Poisson models for insurance ratemaking", *Insurance: Mathematics and Economics*, 48, 226-236.
- Boucher, J.P. and Denuit M., 2008, "Credibility premiums for the zero-inflated Poisson model and new hunger for bonus interpretation", *Insurance: Mathematics and Economics*, 42, 727-735.
- Cameron, A.C. and Trivedi, P.K., 2013, "Regression Analysis of Count Data", *Cambridge University Press*, Second Ed., USA, 189
- Flynn, M. , 2009, "More Flexible GLMs Zero-Inflated Models and Hybrid Models", *Casualty Actuarial Society E-Forum*, 148-224 (https://www.casact.org/pubs/forum/09wforum/flynn_francis.pdf)
- Greene, W., 1994, "Accounting for Excess Zeros and Sample Selection in Poisson and Negative Binomial Regression Models", *NYU Working Paper*, No. EC-94-10 (http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1293115)
- Greene, W., 2007, "Functional Form and Heterogeneity in Models for Count Data", *Foundations and Trends in Econometrics*, 1(2), 113-218.
- Gurmu, S., 1998, "Generalized hurdle count data regression models", *Economic Letters*, 58, 263-268.
- Hardin, J.W., Hilbe J.M., 2012, "Generalized Linear Models and Extensions", Third Ed., Stata Press, Texas.
- Hurlimann W., 1990, "Pseudo Compound Poisson Distributions in Risk Theory", *ASTIN BULLETIN*, 20(1), 57-79.
- Lambert, D., 1992, "Zero-Inflated Poisson Regression, with an Application to Defects in Manufacturing", *Technometrics*, 34(1), 1-17.
- Loeys, T. et.al., 2012, "The analysis of zero- inflated count data: Beyond zero-inflated Poisson regression", *British Journal of Mathematical and Statistical Psychology*, 65, 163-180.
- Morata, L.B., 2009, "A Priori ratemaking using bivariate Poisson regression models", *Insurance: Mathematics and Economics*, 44, 135-141.
- Mouatassim, Y., Ezzahid, E.H., 2012, "Poisson Regression an Zero Inflated Poisson Regression: Application to Private health insurance data", *Eur. Actuar. J.*, DOI 10.1007/s13385-012-0056-2, Springer

Mouatassim, Y., Ezzahid, E.H. and Belasri Y., 2012, "Operational Value at Risk in case of Zero Inflated Frequency", *International Journal of Economics and Finance*, 4(6), 70-77.

Mullahy, J., 1986, "Specification and Testing of some modified count data models", *Journal of Econometrics*, 33, 341-365.

Potts, J.M. and Elith, J., 2006, "Comparing species abundance models", *Ecological Modelling*, 199, 153-163.

Saffari, E.S., Adnan, R. and Greene W., 2012, "Hurdle negative binomial regression model with right censored count data", *SORT*, 36(2), 181-194.

Vuong, Q.H., 1989, "Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses", *Econometrica, JSTOR*, 57(2), 307-333.

Yip, K.C.H. and Yau, K.K.W., 2006, "On modeling claim frequency data in general insurance with extra zeros", *Insurance: Mathematics and Economics*, 36, 153-163.

Xie, H. et.al., 2013, "Comparing statistical methods for analyzing skewed longitudinal count data with many zeros: An example of smoking cessation", *Journal of Substance Abuse Treatment*, 45, 99-108.

Zhao, X. and Zhou X., 2012, "Copula models for insurance claim numbers with excess zeros and time-dependence", *Insurance: Mathematics and Economics*, 50, 191-199.

Zorn, C.J.W., 1996, "Evaluating Zero-Inflated and Hurdle Poisson Specifications", *Midwest Political Science Association*, April 18-20.

Appendix 1: Parameter estimates of models

Parameters	Poisson Model	Negatif Binomial M.	ZIP	ZINB	HURDLE	NBHURDLE
(Intercept)	-1,2848	-1,2852	-0,7048	-1,1671	-0,7509	-1,2656
	< 2e-16***	<2e-16***	<6,4e-10***	2,1e-11***	4,6e-10***	4,6e-07***
genderMALE	0,1112	0,1125	0,1076	0,1262	0,1131	0,1249
	0,0109*	0,0214*	0,2387	0,1050	0,2289	0,2477
maritalstatusMARRIED	0,0036	0,0415	-0,0780	-0,1153	-0,0175	-0,0230
	0,9491	0,9475	0,5296	0,3110	0,8842	0,8675
maritalstatusDIVORCED	0,1972	0,1996	0,2956	0,1845	0,3821	0,4121
	0,0452*	0,0739	0,1354	0,3240	0,0448*	0,0703
maritalstatusWIDOW	0,1378	0,1425	-0,1504	-0,1727	-0,0636	-0,0817
	0,3347	0,3726	0,6634	0,4920	0,8448	0,8235
vehicleage6-10	-0,1767	-0,1774	-0,3156	-0,2143	-0,3655	-0,3874
	0,0006***	0,0018**	0,0066**	0,032*	0,0027**	0,0043**
vehicleage11-15	-0,3209	-0,3217	-0,9987	-0,5339	-0,8408	-0,8896
	0,0868	0,1146	2,3e-07***	0,023*	0,1408	0,1440
vehicleage16-20	-1,0264	-1,0221	-1,6924	-1,2397	-5,6895	-8,3310
	0,3049	0,3266	0,0913	0,2330	0,8686	0,9466
insrdsage36-50	-0,0759	-0,0776	0,0515	0,0861	0,0510	0,0501
	0,1028	0,1392	0,6260	0,3490	0,6060	0,6609
insrdsage51-70	-0,2424	-0,2438	-0,0059	0,0080	-0,0387	-0,0428
	6,4e-06***	0,00005***	0,9611	0,9480	0,7387	0,7482
insrdsage71-100	-0,3843	-0,3873	-0,2942	-0,2365	-0,2020	-0,2233
	0,0017**	0,0041**	0,2842	0,2900	0,4773	0,4831
(Intercept)			-0,2480	-2,5090	-0,2603	-1,2603
			0,2840	0,2300	<2,6e-16***	<2,6e-6***
genderMAN			-0,0125	0,1670	0,1106	0,1106
			0,9470	0,7600	0,0444*	0,0444*
maritalstatusMARRIED			-0,1910	-1,1560	0,0075	0,0075
			0,4660	0,0700	0,9163	0,9163
maritalstatusDIVORCED			0,1790	-0,5040	0,1221	0,1221
			0,6260	0,5200	0,3455	0,3455
maritalstatusWIDOW			-0,7050	-10,8460	0,1942	0,1942
			0,4160	0,9500	0,2771	0,2771
vehicleage6-10			-0,3280	-0,2760	-0,1286	-0,1286
			0,2210	0,6600	0,041*	0,041*
vehicleage11-15			-12,10	-15,6060	-0,2246	-0,2246
			0,9840	1,0000	0,3020	0,3020
vehicleage16-20			-21,10	-15,1410	-0,9111	-0,9111
			1,0000	1,0000	0,3856	0,3856

Appendix 1: Parameter estimates of models (continued)

insrdsage36-50	0,2990	1,7970	-0,1142	-0,1142
	0,1970	0,2900	0,0551	0,0551
insrdsage51-70	0,5350	2,3140	-0,2960	-0,2960
	0,034*	0,2100	1,2e-05***	1,2e-05***
insrdsage71-100	0,2020	1,6150	-0,4336	-0,4336
	0,7410	0,4100	0,0036**	0,0036**

*Significant at the 5% level, ** Significant at the 1% level, ***Significant at the 0,1% level