Available online at http://www.journalcra.com

International Journal of Current Research
Vol. 7, Issue, 10, pp.21198-21204, October, 2015

INTERNATIONAL JOURNAL OF CURRENT RESEARCH

ISSN: 0975-833X

RESEARCH ARTICLE

PROBABILISTIC MODELLING IN THE ANALYSIS OF TRAFFIC CRASH DATA: A STUDY OF BUS DRIVERS' INJURIES IN TURKEY

*,[1]Özge AKKUŞ and [2]Hüseyin TATLIDİL

[1]Department of Statistics, Muğla Sıtkı Koçman University, Turkey
[2]Department of Statistics, Hacettepe University, Turkey

ARTICLE INFO

Article History:
Received 28th July, 2015
Received in revised form
15th August, 2015
Accepted 05th September, 2015
Published online 20th October, 2015

Key words:

Bus accidents, Crash data, Injury severity, Statistical modelling, Probabilistic models.

ABSTRACT

The first aim of this study is to introduce the probability models in modelling the traffic crash data and present comprehensive interpretations of the statistical findings. The second aim is to reveal the risk factors that have important effects on the driver injury severity in bus accidents in Turkey. Data collected by The Department of Traffic Training and Research of General Directorate of Public Security and General Command of Gendarmerie were used. By reference to the ordinal nature of the dependent variable, Ordered Probit Model (OPM) and Ordered Logit Model (OLM) were constructed. The application results show that the average risk of injuries increases along with a 56-year-old or older driver. Other factors leading to the increases in the probability of injury severity are traveling on province road, rainy weather conditions and the accidents occurring as bumping to a stationary material. Accidents occurring as bumping from rear-end or one side of vehicle are is all associated with less severe injuries. Besides the remarkable interpretations of the model results, we also tried to find out which probabilistic model in what conditions fits well in the analysis of a crash data?

Citation: Özge AKKUŞ and Hüseyin TATLIDİL, 2015. "Probabilistic modelling in the analysis of traffic crash data: A study of bus drivers' injuries in Turkey", *International Journal of Current Research*, 7, (10), 21198-21204.

INTRODUCTION

Road transportation is mostly preferred in Turkey compared with the alternatives such as railway, sea or airline. Thus, more frequent traffic accidents are occurred in road transportation. Traffic accidents result in injuries, deaths and economical losses. At this point, statistical information is needed for providing traffic safety, new investments and taking measures. These measures are significantly advanced in parallel with the development of new strategies determined on the basis of the statistical findings. Many models could be used for assessing the risk factors for the traffic crash data but each model requires specific assumptions. Probabilistic models are frequently preferred because researchers could easily obtain the probability values of the severity of injury on condition that the model assumptions are satisfied. Multinomial models could successfully be used if the type of the dependent variable is nominal while ordered models are the best choice when the dependent variable is ordered with J categories. Although dependent variable is actually ordered but we mistakenly use unordered models, we could not obtain efficient parameter estimates. Similarly, if the measure of the dependent variable is actually unordered but we use ordered models, parameter estimates will be biased. When we compare these two types of statistical defectiveness, obtaining biased estimates will cause more serious problems. In our study, we statistically tested and accepted the ordinal nature of the dependent variable. Therefore, we focused on the ordered model results.

****Corresponding author: Özge AKKUŞ***
Department of Statistics, Muğla Sıtkı Koçman University, Turkey.

Model assumptions for some popular probability models are summarized in Table 1 (Akkuş and Özkoç, 2012)

Literature review

Some of the studies in the current literature on traffic crash modelling using probabilistic models are presented below. O'Donnell and Connor (1996) compared the OPM and OLM results by taking the driver attributes as a function of injury severity. In order to examine the risk of different injury levels in all crash types, single-vehicle crashes and two-vehicle crashes in America, Kockelman and Kweon (2002) used OPM with reference to the vehicle characteristics and type of collision. Khattak *et al*. (2002) examined the effects of driver attributes, vehicle factors and type of collision on rolling of vast vehicles and to reveal major factors influencing the severity of injury after the accidents in single-vehicle crashes in North Carolina. They applied Binary Probit Model for rolling tendency and OPM for injury severity.

Table 1. Polychotomous Dependent Variable Models and Assumptions

Model	Dependent Variable Type	Model Assumptions
Multinomial Logit	Nominal	*Only the characteristics of individuals are required. *Strict assumption of Independence of Irrelevant Alternative (IIA) has to be satisfied.
Multinomial Probit	Nominal	*Only the characteristics of indivuduals are required. * No other assumption is necessary including IIA.
Ordered Logit	Ordered	*Only the characteristics of indivuduals are required. * Parallel Slopes Assumtion (PSA) is required.
Ordered Probit	Ordered	*Only the characteristics of indivuduals are required. * Parallel Slopes Assumtion (PSA) is required.
Nested Logit	Nested Nominal Design	*Inclusive Value (IV) are required to be positive.
Conditional Logit	Nominal	*Characteristics of the choice and individuals are both required.
Sequential Lojit	Nested Sequential Design	*Probability of preference in each sequential step is independent from the other probabilities.

*Logit and Probit models only make difference from the link function they used.

In the study of Quddus *et al*. (2002) in Singapore, road conditions, driver characteristics, environmental factors and motor characteristics are determined as a function of driver injuries and severity of motorcycle damages after an accident. According to these characteristics, the severity of injury and the level of vehicle damages were investigated using OPM. Singleton, Qin and Luan (2004) determined factors influencing the injury severity level in occupants of motor vehicles in Kentucky between 2000 and 2001. Uçar and Tatlıdil (2005) compared the results obtained from three probability models for binary outcomes via the data set on motorcycle accidents. Zambon and Hasselberg (2006) studied on the severity of injuries among young motorcyclists in Sweden. Uçar and Tatlıdil (2007) applied OPM in order to determine the factors influencing the severity of damage in bus accidents in Turkey. Savolainen and Mannering (2007) used Nested and Multinomial Logit Models to view the injury severity of motorcyclists in a single and multi-level crash types. Wang and Abdel-Aty (2008) examined the left-turn crash injury severity using Partial Proportional Odds Models. In the study by Milton *et al*. (2008), the severity level of highway was analyzed by the method of Mixed-Logit Model. In their study of Xie *et al*. (2009), risk factors influencing the crash injury severity were analyzed using Bayesian approach. Findings from the Bayesian and classical OPM were compared, as well. Hanrahan, Layde, Zhu, Guse and Hargarten (2009), to quantify the association of driver's age with the risk of being injured, dying, and experiencing injuries of different severity when involved in a motor vehicle crash. Soori, Royanian, Zali and Movahedinejad (2009) studied road traffic injuries in Iran.

Rifaat and Chin (2010) investigated most relevant factors affecting the injury severity by the method of OPM. A comprehensive research on the highway motor-vehicle crash-injury severities was made by Savolainen *et al*. (2011) using some popular discrete choice models. Zhu and Srinivasan (2011) made a study to determine factors influencing the injury severity of large-truck crashes. Mergia *et al*. (2013) modeled the crash injury severity in Ohio using Generalized OLM. Hosseinpour *et al*. (2013) examined the effect of roadway characteristics on the frequency and head-on crashes using the data from Malaysian between the periods of 2007-2010. Kim, Ulfarsson, Kim and Shankar (2013) used mixed logit model in the family of probability models to determine the level of driver injury severity in single-vehicle crashes in California.

In the study of Xi *et al*. (2014), Binary Logit Model is applied to the data recorded from four different regions of China in order to predict the severity in traffic crashes on curved road. Zhao and Khattak (2014) used OPM, Multinomial Logit and Random Parameter Logit models to identify the best model measuring the injury severity of drivers in motor vehicle accidents. Hanrahan, Layde, Zhu, Guse and Hargarten (2009) investigated the association of driver's age with the injury severity in a motor vehicle crash. In the study by Hao and Daniel (2015), driver injury severity related to inclement weather at highway-rail grade crossings in the United States was examined. Zhao and Khattak (2015) studied motor vehicle drivers' injuries in train–motor vehicle crashes.

Studies in the current literature show that probability models are commonly used in modelling the traffic crash data. The most important point is to decide the correct model taking into consideration the dependent variable type and validity of the model assumptions.

Similar to the studies given above, the major contribution of our study is that it reveals the statistical significant factors having an effect on the severity of drivers in bus accidents in Turkey. Additionally, it presents comprehensive interpretations of the ordered model results in terms of the 'Estimated Model Parameters'; 'Estimated Probabilities'; 'Odds Ratios' for OLM and 'Marginal Effect' of an explanatory variable on the estimated probability. It also emphasizes some basic important points in the model selection considering the data structure.

METHODOLOGY

The fundamentals of the ordered response models are based on the continuous underlying (latent) variable that reflects the underlying tendency of observation i on the dependent variable. Ordinality assumption of the dependent variable categories is another important specification of these models. It is assumed that all the explanatory variables are a linear function of the latent variable Y^* in Eq.(1)

$$Y_i^* = \sum_{k=1}^{K} \hat{\beta}_k x_{ik} + \varepsilon_i \; ; \; i = 1, \ldots, n \qquad \ldots\ldots..(1)$$

where X_k $(k = 1, 2, \ldots, K)$ denotes the explanatory variables; $\hat{\beta}$ is the estimated model parameters and ε is a random disturbance term. The random error term ε is assumed to follow Normal or Logistic distribution in ordered response model. If the error term is assumed to follow Normal distribution with mean zero and unit variance, the model is called OPM, where OLM is obtained when the logistic distribution is assumed for the error term with mean zero and variance $\pi^2/3$.

Supposing that the dependent variable is ordered with J categories (where i denotes observation i) and μ_i represents the threshold parameter, the relationship between the observed levels and underlying tendency is given as the following.

$$\begin{aligned} &Y_i = 1, && Y_i^* \le \mu_1 (= 0) \\ &Y_i = 2, && \mu_1 < Y_i^* < \mu_2 \\ &Y_i = 3, && \mu_2 < Y_i^* < \mu_3 \\ &\vdots && \\ &Y_i = J, && \mu_{J-1} < Y_i^* \end{aligned} \qquad \ldots(2)$$

The aim is to obtain the best estimation of an unknown β vector and threshold parameters (μ) using the method of Maximum Likelihood Estimator (MLE). Greene (2000) suggested that the first threshold parameter should be set to zero $(\mu_1 = 0)$. For example, if the dependent variable has J ordered categories, the number of the estimated threshold parameters will be (J-1) in order to constitute J probability area. Because the first threshold parameter is zero, total number of the estimated threshold parameters will decrease to (J-2).

In general, the probability expressions of the ordered models are expressed by using the latent variable approach as follows.

$$P(Y = j) = F\left[\mu_j - \sum_{k=1}^{K} \hat{b}_k x_k\right] - F\left[\mu_{j-1} - \sum_{k=1}^{K} \hat{b}_k x_k\right] \qquad \ldots\ldots\ldots\ldots (3)$$

$$P(Y \le j) = F\left[\mu_j - \sum_{k=1}^{K} \hat{b}_k x_k\right] \qquad \ldots\ldots\ldots\ldots (4)$$

These probabilities will be positive if the threshold parameters satisfy the restrictions of $\mu_1 < \mu_2 < \ldots\ldots < \mu_{J-1}$. (for details, see Liao 1994). The general mathematical forms of the probability expressions for OPM are given by Eq.(5).

$$\begin{aligned} P(Y = 1) &= \Phi\left[-\sum_{k=1}^{K} \hat{b}_k x_k\right] \\ P(Y = 2) &= \Phi\left[\mu_2 - \sum_{k=1}^{K} \hat{b}_k x_k\right] - \Phi\left[-\sum_{k=1}^{K} \hat{b}_k x_k\right] \qquad \ldots\ldots\ldots (5) \\ &\vdots \\ P(Y = J) &= 1 - \Phi\left[\mu_{J-1} - \sum_{k=1}^{K} \hat{b}_k x_k\right] \end{aligned}$$

Φ in Eq.(5) represents the cumulative standard normal distribution function. The following equations introduce the probability expressions of OLM, which are derived on condition that the error term has a logistic distribution denoted by Λ.

$$\begin{aligned} P(Y = 1) &= \Lambda\left[-\sum_{k=1}^{K} \hat{b}_k x_k\right] = \frac{\exp\left(-\sum_{k=1}^{K} \hat{b}_k x_k\right)}{1 + \exp\left(-\sum_{k=1}^{K} \hat{b}_k x_k\right)} \\ P(Y = 2) &= \Lambda\left[\mu_2 - \sum_{k=1}^{K} \hat{b}_k x_k\right] - \Lambda\left[-\sum_{k=1}^{K} \hat{b}_k x_k\right] = \left\{\frac{\exp\left(\mu_2 - \sum_{k=1}^{K} \hat{b}_k x_k\right)}{1 + \exp\left(\mu_2 - \sum_{k=1}^{K} \hat{b}_k x_k\right)}\right\} - \left\{\frac{\exp\left(-\sum_{k=1}^{K} \hat{b}_k x_k\right)}{1 + \exp\left(-\sum_{k=1}^{K} \hat{b}_k x_k\right)}\right\} \\ &\vdots \\ P(Y = J) &= 1 - \Lambda\left[\mu_{J-1} - \sum_{k=1}^{K} \hat{b}_k x_k\right] = 1 - \left\{\frac{\exp\left(\mu_{J-1} - \sum_{k=1}^{K} \hat{b}_k x_k\right)}{1 + \exp\left(\mu_{J-1} - \sum_{k=1}^{K} \hat{b}_k x_k\right)}\right\} \end{aligned} \qquad \ldots\ldots\ldots(6)$$

Model (β) and threshold parameter (μ) vectors could easily be estimated by the method of MLE. The set of best estimators that maximize the likelihood function could be obtained by the classical Newton-Raphson algorithm (Liao 1994; Borooah 2002).

As can be seen in Table 1, ordered probability models are constructed on the validity of the "Parallel Slopes Assumption". It assumes that all the model parameters (β) are the same across the categories of the dependent variable but model equations only differ in terms of the threshold parameters (μ).

In this section, we have given a brief theory of the ordered probability models. Readers who want to learn more about the ordered response and other probability models may turn to Aldrich and Nelson (1984), Liao (1994) and Borooah (2002). Additionally, Table 1 will guide researchers who are not certain in the selection of the most suitable model that fits their data.

Application

We used data collected by General Directorate of Public Security and General Command of Gendarmerie in Turkey. According to the 'Historical Statistics' in Traffic Accident Statistics Reports of Turk Stat between the period of 2003-2013, the rate of bus accidents involved in accidents with respect to registered motor vehicles is at the highest level in Turkey. Thus, we specifically studied the injury severity level of drivers in bus accidents. The total number of recorded bus accidents is 3467. Of these 3467 drivers, the ratios of Uninjured $(Y=0)$; Injured $(Y=1)$ and Death $(Y=2)$ are 89.1%, 10.3% and 0.6%, respectively. Number of deaths in traffic crash data only involves deaths detected at accident scene.

Data

All variables included in the study are given in Table 2.

Table 2. Variables and Descriptions

Dependent Variable (Driver Injury)	
(0) Uninjured (1) Injured	(2) Death
Driver Characteristics	
Age group	(18-25) – (26-35) – (36-45) – (46-55) –(56+)
Education Level	Primary school Secondary school High school Higher education level
Accident Characteristics	
Accident location	Avenue or street – Superhighway – Province road – Other
Number of vehicle(s) in the accident	Single vehicle Two vehicles-same direction Two vehicles-adjacent direction Two vehicles-opposite direction More than two vehicles
Type of collision	Head-on Rear-end (RE); One-side of vehicle (OSV); To stable vehicle (SV); To stationary material (SM) Rollover (R); Other types of collision other than RE, OSV, SV, SM and R
Road Characteristics and Weather Conditions	
Type of covering	Concrete – Asphalt – Parquet – Gravel
Date	Daytime – Night – Twilight
Weather conditions	Clear – Foggy – Rainy – Snowy –Other (cloudy, gusty, snowstormy)
Vehicle Age (year)	

Explanatory variables are composed of driver characteristics including age and education of drivers; accident characteristics including location of the accident, number of vehicles and type of collision; road characteristics including type of covering of the road and day-weather conditions and vehicle characteristic including age of vehicle.

Considering that the vast majority of bus drivers are male, gender of the driver is excluded from the study. Similarly, because accidents occurring outside the settlements are highly correlated with the accidents in state-of-province and vehicle age, accident location is removed from the study in order to eliminate this dependency. Due to the high correlation between the road surface and weather conditions, variable indicating the road surface is also removed from the analysis. Accidents commonly occur with a high ratio in settlement location (79.1%), on avenue or street (66.1%) by single vehicle (33.9%) or two vehicles in the same way (28.4%), on two-way road (50.4), on dry road surface (77.6%), rear-end collision or side collision type (31.8%), day time (72.2%) on outdoors (72.2).

RESULTS

OPM and OLM results are presented in Table 3.

Interpretations were presented according to the signs of the parameters and marginal effects of an explanatory variable on the estimated probabilities in OPM. Different from the OPM results, odds ratio values were also be interpreted in OLM. Nlogit 4.0 package was used in the analyses.

Estimated model and thresholds parameters were presented in Table 3. Before testing the significance of the model parameters, -2LLR (-2xLog Likelihood Ratio) test statistic value was given. This test measures the significance of the statistical difference between the null and estimated model. Results show that both models are statistically significant ($p=0.00<0.05$) and indicate that at least one explanatory variable has an effect on the level of injury of drivers. Additionally, estimated threshold parameters are also significant ($p=0.00<0.05$) in each model and this result justify our prior belief about the ordinal nature of the dependent variable. In other words, the ordinality of the dependent variable is statistically tested and accepted at %5 significance level.

Interpretations of the estimated coefficients

Keeping in mind the reference categories of each variable group, interpretations according to the signs of the coefficients are provided here. Significant and negative coefficient indicates decreases in the severity of injury along with the increases in the associated exploratory variable whereas significant and positive coefficient increases the probability of more severe injuries.

Statistically significant factors leading to the increases in the driver injury severity are examined, we conclude that accidents occurring as bumping to a stationary material (with the coefficient of 1.0343 in OPM and 1.8056 in OLM) cause the highest severe injuries of all other factors. Age group 56+ (0.4188 in OPM and 0.7661 in OLM) and having an accident on province road (0.4215; 0.7711) tends to increase the severity almost in the same rate. Rainy weathers have the lowest effect on the severity of injury of all significant factors that lead to the increases in the severity of injury. In order to see the degrees of effects among all variables that cause increases in the severity of injury, it will be useful to give the following information.

Risk for a driver to be exposed to more severe injuries in the accidents occurring as bumping to a stationary material is 2.5 times higher than the accidents had by a 56-year old or older driver on the province road whereas this ratio is about 5 times compared with the accidents occurring in rainy weather conditions. The remaining variables that have a positive effect on the injury severity are not statistically significant at a significance level of 5%.

When we have a look over other variables that have negative effects on the injury severity, accidents occurring as bumping rear-end of the vehicle (-0.3346; -0.5959) are 1.5 times safer for drivers in comparison with the accidents occurring as bumping from one side (-0.2185; -0.3615) of the vehicle. Remaining coefficients are not statistically significant and have no effect on driver injury severity.

When we examine the results, we see that the interpretations of the coefficients obtained from OLM are approximately the same as OPM. This is because the product of the probit coefficients by 1.8 approximately gives the logit coefficients. Amemiya (1981) suggests that the most appropriate coefficient is 1.6 by reference to his earlier empirical studies. This implies that the sign and the magnitude of the coefficients are relatively equal in each model.

Table 3. OPM and OLM Results

		OPM		OLM		
Reference Category	Indicators	$\hat{\beta}$	p	$\hat{\beta}$	p	Odds Ratios
Avenue or Street	Superhighway	0.3634	0.08	0.6880	0.07	
	Province road	0.4215	0.00*	0.7711	0.00*	2.1621
	Other	0.3165	0.14	0.5886	0.17	
Single Vehicle	Two vehicles-same direction	-0.1239	0.55	-0.3388	0.39	
	Two vehicles-adjacent direction	-0.0743	0.74	-0.2258	0.59	
	Two vehicles-opposite direction	-0.0342	0.88	-0.2240	0.62	
	More than two vehicles	0.1655	0.43	-0.2028	0.61	
Head-on Crash	Rear-end	-0.3346	0.01*	-0.5959	0.01*	0.5511
	One-side	-0.2185	0.04*	-0.3615	0.06	
	Stable vehicle	0.1917	0.40	0.4103	0.31	
	Stationary material	1.0343	0.00*	1.8056	0.00*	6.0835
	Rollover	0.4365	0.10	0.6753	0.17	
	Other	-0.3884	0.07	-0.7873	0.06	
Concrete	Asphalt	0.1593	0.62	0.2735	0.67	
	Parquet	0.3018	0.49	0.5709	0.49	
	Gravel	-5.9032	1.00	-26.952	1.00	
Daytime	Night	0.1042	0.13	0.1877	0.15	
	Twilight	0.0046	0.98	0.0266	0.93	
Clear Weather	Foggy	0.1307	0.53	0.2271	0.56	
	Rainy	0.1956	0.04*	0.3856	0.03*	1.4705
	Snowy	0.2652	0.10	0.5409	0.06	
	Other	0.0475	0.60	0.1298	0.44	
Age Group 18-25	26-35	0.1139	0.40	0.1863	0.46	
	36-45	0.0104	0.94	-0.0301	0.90	
	46-55	0.0919	0.52	0.1223	0.65	
	56+	0.4188	0.03*	0.7661	0.04*	2.0669
Higher Education Level	Primary school	-0.0346	0.90	-0.1163	0.83	
	Secondary school	-0.2822	0.34	-0.5877	0.29	
	High school	-0.1055	0.72	-0.2857	0.60	
	Vehicle age	-0.6563	0.18	-1.1429	0.23	
Threshold Parameter (μ_1)		1.4095	0.00*	3.0809	0.00*	
Constant		-1.351	0.00*	-2.1669	0.02*	
-2Logaritmic Likelihood Ratio Value		230.157	0.00*	224.749	0.00*	

(*)Statistically significant at a significance level of 5%

The unique difference between the estimated coefficients of OLM and OPM arise from the 'one-side' crash type. It is significant in OPM but has no effect on the injury severity in OLM at a critical significance level.

Odds-Ratio interpretations

The Odds-Ratio (OR) values given in the last column of Table 3 can only be obtained from OLM by exponentiating the estimated coefficients of the indicator categories of the statistically significant variables.

OR value for the ***province road*** is 2.1621. This value implies that odds of being exposed to more severe injuries of drivers instead of being uninjured in the accidents occurring on province road is about 2 times higher than the accidents occurring on avenue or street. This result suggests that the level of injury severity of drivers' increases especially in the accidents occurring on province road.

OR value for the "***bumping from rear-end***" of the vehicle is 0.5511. Because this value is smaller than "1", interpreting the inverse of the value (1.8145) will be easier to understand. In this case, we should change the indicator and reference categories of the variable in interpretations. The odds of being exposed to less severe injuries instead of being killed for drivers in the accidents occurring as bumping from rear-end of the vehicle is 1.8145 times higher than the accidents occurring as head-on crashes.

OR value for the "***bumping to a stationery material***" is 6.0835. The odds of being exposed to more severe injuries instead of being uninjured for drivers in the accidents occurring as bumping to a stationery material is about 6 times higher than the head-on crashes

When the results are summarized; we conclude the same findings as the previous section. That is, having an accident as bumping to a stationery material is the most dangerous crash type for drivers while bumping from rear-end of the vehicle is the safest crash type.

The OR value for the "***rainy weather***" is 1.4705. The odds of being exposed to more severe injuries of drivers instead of being uninjured in the accidents occurring in rainy weathers is 1.4705 times higher than the accidents occurring in clear weather conditions. When the slippery surface of the road is taken into consideration in rainy weathers, it is an expected result to encounter with more severe injuries.

OR value for the "***age group of 56+***" is 2.0669. The odds of being exposed to more severe injuries instead of being uninjured for a 56-year old or older driver is about 2 times higher than the age group 18-25. It is also not surprising because older drivers have different kinds of health problems such as weakness of reflex, having low resistance of their body to the injuries and having a problem with their sight or hearing. In the following section, results will be interpreted in terms of the marginal effects of an explanatory variable on the estimated injury probability.

Because the estimated coefficients of the 'province road', 'bumping to a stationery material', 'rainy weather conditions' and 'age group 56+' are positive, we observe that increases in any of these variables will lead to the marginal increases in the probability of death. Only the accidents bumping to a stationery material (0.0092) substantially raise the probability of death. It also leads to the maximal decreases in the probability of non-injured (-0.1664). Accidents occurring in the province road (0.0038) causes the second largest increases in the probability of death, to be in age group of 56+ (0.0037) and rainy weather conditions (0.0017) lead to smaller increases in the probability of death in comparison with the accidents occurring as bumping to a stationery material in province road.

Since all the estimated parameters of the remaining variables (bumping from rear-end and bumping from one side of the vehicle) are negative, they have increasing effect on the probability of no injury. When we compare these variables, bumping from rear-end leads to larger increases in the probability of no injury (0.0538) than the accidents occurring as bumping from one side of the vehicle (0.0352).

Table 4. Marginal effects of the significant explanatory variables on the estimated probabilities

OPM RESULTS			
Variable	P (Y=Uninjured)	P (Y=Injured)	P (Y=Death)
Province road	-0.0678	0.0641	0.0038
Bumping from rear-end	0.0538	-0.0509	-0.0030
Bumping from one side	0.0352	-0.0332	-0.0019
Bumping to a stationery material	-0.1664	0.1572	0.0092
Rainy weather	-0.0315	0.0297	0.0017
Age group 56+	-0.0674	0.0637	0.0037
OLM RESULTS			
Variable	P (Y=Uninjured)	P (Y=Injured)	P (Y=Death)
Province road	-0.0582	0.0551	0.0031
Bumping from rear-end	0.0450	-0.0426	-0.0024
Bumping to a stationery material	-0.1363	0.1290	0.0074
Rainy weather	-0.0291	0.0275	0.0016
Age group 56+	-0.0548	0.0519	0.0030

Marginal Effects

In previous sub-sections, we have only interpreted the signs of the coefficients, which only indicate the direction of changes in probabilities and odds ratio values. We will also give the magnitudes of these changes via marginal effects associated with each statistically significant coefficient in Table 4.

Because the marginal effects of OPM are approximately the same as OLM, we will only focus on the interpretation of the OPM results.

Marginal effects give the effect of the concerned explanatory variable on the estimated probability while other variables are set to their own mean values. The most important point to be taken into consideration while we interpret the results of the marginal effects is the reference categories of all the variable groups.

Conclusion

We have practically discussed the selection of the most appropriate probability model as considering the prosperity of interpretation even though we know that the broader research literature has shown there is no basis for preferring one framework to the others as both model provide similar results under a wide range of settings. The major problem is the indefiniteness of the basic criteria providing us to justify our choice, theoretically. The unique criteria revealing the superiority of OLM is the odds-ratio interpretations.

As for the application results, some remarkable findings are summarized below.

OLM and OPM results are consistent with each other except from the accident type occurring as bumping from one side of the vehicle. Province road, bumping to a stationary material, rainy weathers and being in age group of 56+ lead to increases in the severity of injury whereas accidents occurring as bumping from rear-end or one side of the vehicle lead to the decreases in the severity of injury. Of all type of collisions, accidents that occur as bumping from rear-end of the vehicle are the safest ones whereas accidents occurring as bumping to a stationary material are the most dangerous crash type for drivers.

REFERENCES

Akkuş, Ö., Özkoç, H. 2012. A comparison of the models over the data on the interest level in politics in Turkey and countries that are members of the European Union: Multinomial or ordered logit model?. *Res J Appl Sci Eng Technol*., 4(19):3646-3657.

Aldrich, JH., Nelson, FD. 1984. Linear Probability, Logit and Probit Models. Sage Publications, Inc: London, 07-045.

Amemiya, T. 1981. Qualitative response models: A survey. J. Econ Literature, 19:1483-1536.

Borooah, VK. 2002. Logit and Probit (Ordered and Multinomial Models. Sage University Papers, 07-138, Inc: London; 2002.

Greene, WH. 2000 Econometric Analysis. New York University, Prince Hall, Upper Saddle River, New Jersey 07458, ISBN: 0-13-013297-7.

Hanrahan, RB., Layde, PM, Zhu, S, Guse, CE, Hargarten, SW. 2009. The association of driver age with traffic injury severity in Wisconsin. *Traffic Injury Prevention*, 10(4): 361-7.

Hao, W., Daniel, J. 2015. Driver injury severity related to inclement weather at highway-rail grade crossings in the United States. Traffic Injury Prevention, DOI: 10.1080/15389588.2015.1034274.

Hosseinpour, M., Yahaya, A., Sadullah, A. 2014. Exploring the effects of roadway characteristics on the frequency and severity of head-on crashes: Case studies from Malaysian federal roads. Accident Anal Prev. 10.1016/j.aap.2013.10.001:209-222.

Khattak, AJ, Schneider, RJ, Targa, F. 2002. Risk factors in large truck rollovers and injury severity: Analysis of single-vehicle collisions. Paper presented at: Transportation Research Board 82nd Annual Meeting, Washington, D.C., USA.

Kim, JK, Ulfarsson, GF, Kim, S, Shankar, VN. 2013. Driver-injury severity in single-vehicle crashes in California: a mixed logit analysis of heterogeneity due to age and gender. Accident Anal Prev., 50: 1073-1081.

Kockelman, KM, Kweon, YJ. 2002. Driver injury severity: An application of ordered probit models. Accident Anal Prev., 34 (3):313-321.

Liao, TF. 1994. Interpreting probability models (logit, probit and other generalized linear models. Sage Publications, Thousand Oaks, Inc: London, 07-101.

Mergia, W., Eustace, D., Chimba, D., Qumsiyeh, M. 2013. Exploring factors contributing to injury severity at freeway merging and diverging locations in Ohio. Accident Anal Prev., 10.1016/j.aap.2013.03.008:202-210.

Milton, JC., Shankar, VN., Mannering, FL. 2008. Highway accident severities and the mixed logit model: An exploratory empirical analysis. *Accident Anal Prev*., 40(1): 260-266.

O'Donnell, CJ., Connor, DH. 1996. Predicting the severity of motor vehicle accident injuries using models of ordered multiple choice, *Accident Anal Prev*., 28(6): 739-753.

Quddus, MA., Noland, RB., Chin, HC. 2002. An analysis of motorcycle injury and vehicle damage severity using ordered probit models, *J Safety Res*., 33 (4):445-462.

Rifaat, SM., Chin, HC. 2010. Accident severity analysis using ordered probit model. J Adv Transport., 41(1):91-114.

Savolainen, P., Mannering, F. 2007. Probabilistic models of motorcyclists' injury severities in single- and multi-vehicle crashes, Accident Anal Prev. 39(5):955-963.

Savolainen, P., Mannering, F., Lord, D., Quddus, MA. 2011. The statistical analysis of highway crash-injury severities: A review and assessment of methodological alternatives. Accident Anal Prev., 43(5):1666-1676.

Singleton, M., Qin H, Luan J. 2004. Factors associated with higher levels of injury severity in occupants of motor vehicles that were severely damaged in traffic crashes in Kentucky, 2000-2001, Traffic Inj Prev., 5(2):144-50.

Soori, H., Royanian, M., Zali, AR., Movahedinejad, A. 2009. Road traffic injuries in Iran: the role of interventions implemented by traffic police, Traffic Inj Prev., 10(4): 378-8.

Turkish Statistical Institute. Traffic Accident Statistics (Road). Available at: http://www.tuik.gov.tr/Kitap.do?metod=KitapDetay&KT_ID=15&KITAP_ID=70, 2013.

Uçar, Ö., Tatlıdil, H. 2005. Application of three discrete choice models to motorcycle accidents and a comparison of the results. Hacettepe J Math Stat., 34: 55-66

Uçar, Ö., Tatlıdil, H. 2007. Factors influencing the severity of damage in bus accidents in Turkey during 2002: An application of the ordered probit model. *Hacettepe J Math Stat*., 36: 79-87.

Wang, X., Abdel-Aty, M. 2008. Analysis of left-turn crash injury severity by conflicting pattern using partial proportional odds models, *Accident Anal Prev*., 40(5):1674-1682.

Xi, J., Liu, H., Cheng, W., Zhao, Z., Ding, T. 2014. The model of severity prediction of traffic crash on the curve. Math Probl Eng., 10.1155/2014/832723:1-5.

Xie, Y., Zhang, Y., Liang, F. 2009. Crash injury severity analysis using bayesian ordered probit models. *J. Transp Eng*., 135(1):18–25.

Zambon, F., Hasselberg, M. 2006. Factors affecting the severity of injuries among young motorcyclists-a Swedish nationwide cohort study. *Traffic Inj Prev*., 7(2):143-149.

Zhao, S., Khattak, A. 2015. Motor vehicle drivers' injuries in train–motor vehicle crashes. Accident Anal Prev., 10.1016/j.aap.2014.10.022:162-168. *Online publication* date: 1-Jan-2015.

Zhu, X., Srinivasan, S. 2011. A comprehensive analysis of factors influencing the injury severity of large-truck crashes. *Accident Anal Prev.,* 43(1): 49-57.

****