KSCE Journal of Civil Engineering (2012) 16(7):1280-1288
DOI 10.1007/s12205-012-1527-3

Transportation Engineering
www.springer.com/12205

Analysis of Injury Severity in Traffic Crashes: A Case Study of Korean Expressways

Sunghee Park*, Kitae Jang, Shin Hyoung Park***, Dong-Kyu Kim****, and Kyung Soo Chon*******

Received March 29, 2011/Revised 1st: July 6, 2011, 2nd: December 26, 2011/Accepted January 18, 2012

Abstract

The objective of this study is to evaluate a set of variables that contribute to the degree of injury severity sustained in traffic crashes of Korean expressways. To this end, we examined three statistical models − ordered probit, ordered logit, and multinomial logit − to determine the most appropriate model for crash records that were collected from the entire network of Korean expressways in 2008. Interpretation of the estimated coefficients in the selected model provides relative risks of significant influential factors for injury severity. The findings from this study are expected to help transportation planners and engineers understand which risk factors contribute more to the injury severity in Korean expressways such that they can efficiently allocate resources and effectively implement safety countermeasures.

Keywords: *traffic crashes, Korean expressways, injury severity, ordered probit, ordered logit, multinomial logit, influential factors*

1. Introduction

Traffic crashes rank sixth among the causes of deaths in Korea, and they are the second-leading cause of the deaths of people between the ages of 1 and 19 (Statistics Korea, 2010). For the recent five-year period from 2005 to 2009, 30,577 people were killed and 1,719,205 were injured on Korea's roadways (Korea Transport Database, 2011). Due to recent policy and planning efforts for enhancing traffic safety by the Korean government, traffic fatalities and injuries tended to decrease every year. Yet, many improvements can still be made, and one major area is attenuating the severity of the injuries sustained on expressways. Although only 6.9% of the fatalities and 2.8% of the injuries occurred on expressways, crashes on expressways, compared to those for other road classifications, have significantly higher ratios of fatalities and injuries to crash occurrences, i.e., 0.11 fatalities and 2.56 injuries per crash, versus 0.027 fatalities and 1.57 injuries per crash for other road classifications. These higher ratios of fatalities and injuries to crash occurrences indicate that the crashes on expressways resulted in more deaths and injuries than crashes on other road classifications. This fact suggests that one of the fundamental approaches for enhancing safety on the expressways is to identify and evaluate the major factors that contribute to the severity of injuries sustained by drivers involved in expressway crashes. To shed light on this issue, the present study evaluates risk factors on aggravating injury severity of traffic crashes in Korean expressways.

The remainder of this paper is organized as follows: The pertinent previous research is reviewed in Section 2, and detailed descriptions of crash data from Korean expressways are provided in Section 3. Section 4 examines three different statistical models to determine the most appropriate model for the collected crash data. Findings from the interpretation of estimated coefficients are documented in Section 5. Section 6 summarizes findings and discusses implications of this study.

2. Literature Review

Evaluation of risk factors for the severity of injuries sustained in traffic crashes has been a major and an essential topic for traffic safety research. Due to its importance, there has been extensive research utilizing various statistical models to unveil the relationship between risk factors and injury severity. This section reviews risk factors reported in previous research (section 2.1), and examines statistical models whether they could be used to assess injury severity involved with traffic crashes in

*Ph.D. Student, Dept. of Civil and Environmental Engineering, Seoul National University, Seoul 151-744, Korea (E-mail: sunny83@snu.ac.kr)
**Assistant Professor, The Cho Chun Shik Graduate School for Green Transportation, Korea Advanced Institute of Science and Technology, Daejeon 305-701, Korea (Corresponding Author, E-mail: kitae.jang@kaist.ac.kr)
***Member, Associate Researcher, Transportation Research Division, Expressway & Transportation Research Institute, Korea Expressway Corporation, Hwaseong 445-812, Korea (E-mail: shinhpark@ex.co.kr)
****Member, BK21 Professor, Dept. of Civil and Environmental Engineering, Seoul National University, Seoul 151-744, Korea (E-mail: kimdk95@snu.ac.kr)
*****Member, Professor, Dept. of Civil and Environmental Engineering, Seoul National University, Seoul 151-744, Korea (E-mail: chonks@snu.ac.kr)

Korean expressways (section 2.2).

2.1 Risk Factors

The risk factors reported in previous research are shown to vary across studies, implying that a set of significant influential factors may be dependent both on the source of crash data and the scope of study (see Table 1). These findings suggest that, in investigating the relationship between risk factors and injury severity of traffic crashes, one should not rely purely on the previous research in evaluating a set of influential factors. Therefore, this section furnishes findings in previous research that specifically concerned crash data collected from Korean expressways.

As growing concerns on traffic safety in Korean society, there have been extensive research efforts investigating factors for traffic safety in Korean expressways (e.g., Lee *et al*., 2007; Oh *et al*., 2010; Park *et al*., 2009; Ryu *et al*., 2006). Majority of these studies focus on investigating how environmental and traffic attributes affect crash occurrences (as measured by crash count per prescribed unit such as roadway segment, intersection, etc.). Meanwhile, studies regarding risk factors increasing injury severity of traffic crashes in Korean expressways are rare, despite of their importance for understanding crash outcomes. Only one recent study (Lee *et al*., 2011) has been found which used crash data collected from a segment of Korean expressways (sample size of 489), limiting to two lower-levels of injury severity. Hence, the findings from this study may not represent the factors contributing to the injury severity of traffic crashes in the entire network of Korean expressways. Furthermore, there could be possible bias toward lower-level injuries by excluding samples with high-level of injury severity. Therefore, further research using a more comprehensive and abundant data sources is needed to provide more general and unbiased evidence. In view of this, the study herein examines all available factors in crash database from scratch in an aim to select the set of variables that are the most suitable for crash data in Korean expressways.

Table 1. Summary of the Pertinent Literature

Author	Objective	Method	Influential Factors
Lee *et al.* (2008)	factors causing vehicle crashes	ordered probit model	· winter · major road left-turn lanes · industrial region · high ADT of major road
Park *et al.* (2008)	the characteristics of accidents and the accident factors that affect severity	ordered probit model	· minor road traffic volumes · major road left-turn lanes · major road yellow signal time · major and minor road speed limit
Singleton *et al.* (2004)	risk factors on injury severity	ordinal logistic regression with stepwise selection	· older driver · female driver · not wearing seat belt · ejection from vehicle · alcohol use · rollover · fire · head-on collision · collision with fixed object · federal/state road way · speed
Abdel-Aty (2003)	influential factors on level of driver injury	ordered probit model, multinomial logit model, nested logit model	· female driver · older driver · not wearing seat belt · speed ratio · driver struck at her/his side · passenger car
Kockelman and Kweon (2002)	risk factors on driver injury severity	ordered probit model	· older driver · older vehicle · alcohol · rollover · female driver · nighttime
O'Donnell and Connor (1996)	influential factors on the severity of traffic crash injuries	ordered probit model/logit model	· age of victims · vehicle speed · seating position · blood alcohol level · vehicle type · vehicle make · type of collision

2.2 Model Specification

The severity of injuries measured for crash records has both categorical and ordinal characteristics (e.g., KABCO and Maximum Abbreviated Injury Scale (MAIS)[1)]). Hence, many previous studies have used models with ordered structure to analyze risk factors and their effects on the severity of injuries sustained in traffic crashes (see Table 1).

On the other hand, Washington *et al*. (2003) addressed a potential issue that an ordered probability model may be inappropriate due to the restrictions on how variables affect ordered discrete outcome probabilities even though the data are ordinal. The authors suggested that, since a trade–off is inherently being made between recognizing the ordering of responses and losing the flexibility in specification offered by unordered probability models, one should pay attention to select ordered and unordered models.

Although ordered models may be plagued by such ambivalent effects, only a few studies have compared and examined the performance of statistical models. Abdel-Aty (2003) and Abdel-Aty and Abdelwahab (2004) analyzed the severity of drivers' injuries. Those studies used an ordered probit model, a multinomial logit model and a nested logit model, and determined the ordered probit model to be the best model because of its accuracy and simplicity of use. Since this finding is based on limited empirical evidence, it requires further confirmation because it is likely that the most appropriate model specification may be dependent on the data, as occurred in selecting the set of risk factors. Therefore, a statistical model should be selected carefully to appropriately represent the relation between the severity of injuries and influential factors.

1) Observational injury evaluation scale. KABCO: K=Killed; A=Incapacitating Injury; B=Non-Incapacitating Injury; C=Possible Injury; O=No Injury, and MAIS: MAIS 1 = Minor Injury; MAIS 2 = Moderate Injury; MAIS 3 = Serious Injury; MAIS 4 = Severe Injury; MAIS 5 = Critical Injury; Fatal (Tarko *et al*., 2010).

3. Data Description

To analyze influential factors on the severity of injuries sustained in traffic crashes, all crash records from the entire network of Korean expressways (excluding privately invested expressways) in 2008 were used. These data were provided by the Korea Expressway Corporation. The database contains 53 independent variables and one dependent variable. The dependent

Table 2. Variables and Descriptive Statistics

Variable categories	Variables	Percentage (%)	Number of crashes			
			SUM	Fatality	Injury	Property damage only
Dependent variable						
Level of injury severity	Fatality	2.3	230	-	-	-
	Injury	9.4	947			
	Property damage only	88.3	8,862			
Independent variables						
Time	Midnight - 6:00 A.M.	18.1	1,814	73	221	1,520
	6:00 A.M. - Noon	28.5	2,864	40	216	2,608
	Noon - Midnight	53.4	5,361	117	510	4,734
Emergency service arrival time	- 15min	62.1	6,235	150	605	5,480
	15 - 20 min	15.5	1,558	44	153	1,361
	20 - 25 min	11.8	1,187	21	104	1,062
	25 - 30 min	5.8	575	11	54	510
	More than 30 min	4.8	484	4	31	449
Crash location	Main road	69.3	6,961	200	770	5,991
	Ramp	12.3	1,236	10	69	1,157
	Toll gate	12.9	1,295	3	47	1,245
	Others	5.5	547	17	61	469
Primary crash factor	Vehicle defects*	8.9	895	15	55	825
	Obstacles, poor road conditions	11.5	1,159	0	16	1,143
	Alcohol use	1.7	173	6	28	139
	Speeding	18.3	1,833	36	138	1,659
	Driving while drowsy	16.0	1,602	73	266	1,263
	Lack of visual attention	13.5	1,356	37	160	1,159
	Median violation	0.2	18	5	8	5
	Excessive steering	14.5	1,453	19	130	1,304
	Others	15.4	1,550	39	146	1,365
Weather	Rainy or snowy	18.6	1,870	36	147	1,687
	Clear	62.6	6,281	152	626	5,503
	Others (cloudy, windy, misty)	18.8	1,888	42	174	1,672
Collision partner	Fire	2.2	223	5	6	212
	Car vs. car	12.8	1,286	98	337	851
	Car vs. facility	53.8	5,395	54	422	4,919
	Car vs. people	0.6	63	33	21	9
	Car only	19.1	1,919	38	137	1,744
	Others	11.5	1,153	2	24	1,127
No. of parties	1	75.8	7,609	90	454	7,065
	2	18.4	1,846	107	371	1,368
	3	3.7	373	16	81	276
	More than 3	2.1	211	17	41	153
Radius of curvature	Straight	64.7	6,500	164	616	5,720
	Right curve more than 500m	13.5	1,354	20	130	1,204
	Left curve more than 500m	13.7	1,380	35	142	1,203
	Right curve less than 500m	4.4	437	3	30	404
	Left curve less than 500m	3.7	368	8	29	331
Vertical grade	Downhill	29.8	2,987	75	307	2,605
	Flat	46.8	4,701	107	371	4,223
	Uphill	23.4	2,351	48	269	2,034
type of vehicle at fault	Passenger car	61.1	6,136	121	544	5,471
	Van	8.7	874	31	102	741
	Truck, trailer	29.0	2,904	76	290	2,538
	Others	1.2	125	2	11	112
Driver age	Less than 20	0.2	25	2	2	21
	20 - 29	15.7	1,572	39	176	1,357
	30 - 39	23.8	2,392	49	234	2,109
	40 - 49	26.3	2,640	72	240	2,328
	50 - 59	16.0	1,602	42	171	1,389
	Equal to or more than 60	3.7	375	11	41	323
	Unknown	14.3	1,433	15	83	1,335

*The vehicle defects variable includes various defections of physical devices such as power system, engine, electric system, steering, brake system and tire.

variable measured the severity of injuries, which is categorized into three levels: 1) property damage only (PDO), 2) injury, and 3) fatality[2]. The order of injury severity is fatality, injury and PDO. PDO shows the highest proportion, which is 88.3% of all crash data, and injury and fatality comprise 9.4% and 2.3%, respectively.

4. Methods

In this section, the specifications of the model used in this study are described. Since the dependent variable representing levels of injury severity, i.e., property damage only, injury, and fatality, is discrete, three common statistical models for this type of outcome variable were considered, i.e., 1) ordered probit; 2) ordered logit; and 3) multinomial logit. The ordered probit and logit models assume that the severity of injuries has an ordinal nature, while the multinomial logit does not make this assumption. By comparing these two groups of models (i.e., the one that assumes ordinal nature and the other that does not), the effect of the assumption of ordinal nature on the outcome variable can be examined. Explicit specifications of these models are furnished below.

4.1 Ordered Probit and Logit Models

The ordered models can be specified as follows:

$$\boldsymbol{I}_p^* = \boldsymbol{\beta}'\boldsymbol{X}_p + \varepsilon_p \tag{1}$$

where, $\boldsymbol{I}_p^*$ is an unobserved continuous descriptor variable that measures the severity of the injury in the p^{th} crash; $\boldsymbol{\beta}$ is a vector of unknown parameters; $\boldsymbol{X}_p$ is a vector of the measurable and observed variables that describe the characteristics of the p^{th} crash (e.g., time, emergency arrival time, and crash location); and ε_p is an error term of the p^{th} crash.

The only difference between the ordered probit and logit models is the assumption for the distribution of the error term i.e., the error term, ε_p, in the ordered probit model is specified to follow a standard normal distribution, while this term in the logit model is specified to follow a logistic distribution.

We can only observe $\boldsymbol{I}_p$ (injury categories) from a given dataset, because $\boldsymbol{I}_p^*$ is an unobservable, underlying, continuous descriptor. Thus, Eq. (1) for $\boldsymbol{I}_p^*$ can be translated into a form of censoring:

$$\boldsymbol{I}_p = \begin{cases} 1 \text{ if } -\infty < \boldsymbol{I}_p^* \le \psi_1 \text{ PDO} \\ 2 \text{ if } \psi_1 < \boldsymbol{I}_p^* \le \psi_2 \text{ Injury} \\ 3 \text{ if } \psi_2 < \boldsymbol{I}_p^* \le \infty \text{ Fatality} \end{cases} \tag{2}$$

where, the threshold values (ψ_1, ψ_2 and $\psi_1 \le \psi_2$) are unknown parameters to be estimated along with $\boldsymbol{\beta}$. (For more information, see Jang *et al*. (2010) and McKelvey and Zavoina (1975)). Then, the probability of:

$$\mathrm{P}(\boldsymbol{I}_p = i) = \mathrm{F}(\boldsymbol{I}_p^* < \psi_i - \boldsymbol{\beta}'\boldsymbol{X}_p) - \mathrm{F}(\boldsymbol{I}_p^* \le \psi_{i-1} - \boldsymbol{\beta}'\boldsymbol{X}_p) \tag{3}$$

where, $\mathrm{P}(\boldsymbol{I}_p = i)$ is the probability that the p^{th} crash results in i level of injury (i = 1, 2, and 3); $\mathrm{F}(\boldsymbol{I}_p^* < \psi_0 - \boldsymbol{\beta}'\boldsymbol{X}_p) = 0$; $\mathrm{F}(\boldsymbol{I}_p^* < \psi_3 - \boldsymbol{\beta}'\boldsymbol{X}_p)$ =1; and F( ) is the cumulative distribution function (if logit, it is logistic; if probit, it is normal).

4.2 Multinomial Logit Model

The multinomial logit model is usually appropriate to use when the dependent variable is nominal (not ordered in any systematic way). The difference between multinomial and ordered logit models is that the error correlations are set to be zero in the multinomial logit model, while the errors in ordered logit are perfectly correlated (For more information, see Borooah (2001).)

$$\mathrm{P}(\boldsymbol{I}_p = i) = \frac{e^{\beta_i' X_p}}{1 + \sum_i e^{\beta_i' X_p}}, \quad (i = 1, 2, \text{ and } 3) \tag{4}$$

where, $\mathrm{P}(\boldsymbol{I}_p = i)$ is the probability that the p^{th} crash results in i level of injury (i = 1, 2, and 3).

5. Estimated Result

5.1 Model Estimates and Selection

The models specified in Section 4 can be estimated by using commercially available softwares (STATA, SAS, SPSS, etc.). In this study, STATA ver.10 was used for estimation. 11 variables with the highest percentage in each of 11 variable categories are set as default variables, and, thus, the results of 42 out of the 53 independent variables are identified[3]. The p-values are also calculated by using robust standard errors to draw valid interpretations for each variable within 5% level of significance. The model estimates, p-values, marginal effects, and log-likelihood ratio indices of ordered probit, ordered logit, and multinomial logit models are provided in Tables 3, 4, and 5, respectively.

The coefficient of each variable represents how it affects the severity of injury. The positive sign of a coefficient means that the corresponding variable tends to increase the severity level compared to the default variable. On the other hand, coefficients that have negative signs tend to decrease the severity level of injury as the corresponding variable increases.

As mentioned by Spiegel and Stephens (1982), the coefficient gives the relationship between a vector of explanatory variables and the observed dependent variable, and an *unobservable continuous descriptor variable*, I^*, that measures the severity of the injury, as described in Section 4. Therefore, the coefficients cannot be compared across different models. On the other hand, the marginal effects indicate the independent effects of explanatory variables on the changes in the probability of having a certain level of injury. Therefore, they present the effects of explanatory

2) Any occupant involved in a crash died within 30 days from the time of crash.

3) The variables of the highest percentage were selected i) to compare the remaining variables to the highest percentage variable, which is the most common within the category; and ii) to avoid instability and bias due to statistical fluctuations because variables with small samples are be influenced more by (potentially existing) outliers.

Table 3. Model Estimates of Ordered Probit

Variables	OP				
	Coefficient	P>|z|	Marginal effect		
			Fatality	Injury	Property damage only
Time					
Midnight - 6:00A.M.	0.1044088	0.021	0.00259	0.01378	-0.01637
6:00A.M. - Noon	-0.147584	0.001	-0.0031	-0.018	0.021137
Crash location					
Ramp	-0.280803	0	-0.005	-0.03114	0.036182
Toll gate	-0.57782	0	-0.0084	-0.05562	0.063971
Primary crash factor					
Vehicle defects	-0.239322	0.006	-0.0044	-0.0268	0.03115
Obstacles, poor road conditions	-0.991118	0	-0.0108	-0.077	0.087842
Driving while drowsy	0.1512224	0.018	0.00392	0.02039	-0.02431
Median violation	1.099347	0	0.08944	0.2111	-0.30054
Weather					
Rainy or snowy	-0.124818	0.021	-0.0026	-0.01508	0.017681
Collision partner					
Car vs. car	0.4454816	0	0.01541	0.06788	-0.08328
Car vs. people	2.33661	0	0.46415	0.27895	-0.7431
Car only	0.1781157	0.001	0.00466	0.02412	-0.02878
No. of parties					
2	0.4563872	0	0.01508	0.06805	-0.08313
3	0.5409395	0	0.02275	0.08908	-0.11182
More than 3	0.7538032	0	0.04106	0.13444	-0.1755
Type of vehicle at fault					
Van	0.2866939	0	0.00876	0.04164	-0.0504

Threshold value		Coefficient	Standard error
ψ_1 (between PDO and Injury)		1.279538	0.0734303
ψ_2 (between Injury and Fatality)		2.270861	0.078424
Classification	Number of Independent variables	Log likelihood	Adjusted Likelihood Ratio Index
LL(0)	0	-4,209.4555	-
LL(β)	42	-3,540.8524	0.148856093

*This table shows only the variables that satisfy the 5 % significance level.

variables on the probability of level of injury, I, which is observable. For the same reason (also because the same set of variables were used as inputs), the marginal effect is suitable for comparing the effect of variables across models, which is one of the objectives in the present study.

In the ordered probit model, 16 independent variables (except "others" and "unknown" variables) were significant at 5% level. In the ordered logit model, 16 variables were identified as significant, while 17 and 14 variables for fatality and injury level, respectively, were significant in the multinomial logit model.

As shown in Tables 3 and 4, the results of the ordered probit model are comparable to those of the ordered logit model. 16 common variables were chosen as the influential factors, and the signs of the marginal effect for each of severity levels were identical.

On the other hand, the results provided by the multinomial logit model showed considerable differences, as shown in Table 5. Since the multinomial logit model measures the substantive effects of fatality and injury with respect to PDO (not reflecting the perfectly correlated order between injury levels as in the ordered logit and probit models), the p-values and marginal effects of this model are calculated separately for fatality and injury. As a result, only 10 independent variables are significant at the 5% level for both severity levels, and 11 variables, except the "others" and "unknown" variables, meet the 5% significance level for only one severity level.

In all three models, the commonly selected variables have the same sign for each coefficient. This shows that the signs of marginal effects in the multinomial logit model are comparable to the results of the ordered logit and probit model. Also, the order among the marginal effects of PDO, injury, and fatality is identified as the same as the orders among the chosen influential factors of all models, except for one variable, i.e., car vs. people. It can be interpreted that a crash between a car and a person may lead to severe injury or death since drivers on expressways are driving at a high rate of speed, and they do not expect to encounter pedestrians. This overall conformity in the sequence of injury severity between ordered and multinomial models satisfies a necessary (but not sufficient) condition for confirming ordinal nature of injury severity.

To select a model that has the best fit, the log likelihoods of the three models were calculated. Since likelihood ratio index (ρ^2) tends to increase when additional independent variables are entered in the model, the adjusted likelihood ratio index ($\bar{\rho}^2$) is

Table 4. Model Estimates of Ordered Logit

Variables	OP				
	Coefficient	P>|z|	Marginal effect		
			Fatality	Injury	Property damage only
Time					
Midnight - 6:00A.M.	0.208109	0.015	0.002412	0.012915	-0.01533
6:00A.M. - Noon	-0.28271	0.001	-0.0029	-0.01581	0.018709
Crash location					
Ramp	-0.52722	0	-0.00476	-0.02637	0.031126
Toll gate	-1.15203	0	-0.00865	-0.04862	0.05727
Primary crash factor					
Vehicle defects	-0.41744	0.017	-0.00385	-0.02131	0.025166
Obstacles, poor road conditions	-2.0463	0	-0.01203	-0.06822	0.08025
Driving while drowsy	0.286147	0.019	0.003423	0.018223	-0.02165
Median violation	1.97234	0	0.062626	0.230493	-0.29312
Weather					
Rainy or snowy	-0.25159	0.017	-0.00253	-0.01385	0.016384
Collision partner					
Car vs. car	0.772362	0	0.011333	0.0576	-0.06893
Car vs. people	4.276451	0	0.426515	0.350852	-0.77737
Car only	0.317379	0.002	0.003802	0.020229	-0.02403
No. of parties					
2	0.892039	0	0.013121	0.066461	-0.07958
3	1.105214	0	0.020601	0.097752	-0.11835
More than 3	1.459595	0	0.033588	0.14675	-0.18034
Type of vehicle at fault					
Van	0.495628	0	0.006624	0.034481	-0.0411

Threshold value		Coefficient	Standard error
ψ_1 (between PDO and Injury)		2.257398	0.1435797
ψ_2 (between Injury and Fatality)		4.250343	0.1583044
Classification	Number of Independent variables	Log likelihood	Adjusted Likelihood Ratio Index
LL(0)	0	-4,209.4555	-
LL(β)	42	-3,542.1675	0.148543677

*This table shows only the variables that satisfy the 5 % significance level.

used to consider the change in the degrees of freedom. This test is useful to assess two competing models. For more information, see Washington *et al*. (2003). The results of the log likelihood and the adjusted likelihood ratio index are also presented in Tables 3, 4, and 5.

In Tables 3, 4, and 5, LL(0) is the initial log likelihood with all parameters set to zero, while LL(β) is the log likelihood at convergence with a parameter vector. Higher values of the adjusted log likelihood ratio index for a certain model indicate that the model explains the data better than other models. In this study, adjusted log likelihood ratio indices of three models are similar, which suggests that three models have similar goodness of fit to the crash data from Korean expressways.

As discussed above, all three models exhibit similar outcomes and thus can be selected to evaluate risk factors on injury severity caused by traffic crashes in Korean expressways. However, the outcomes from the multinomial model indicate that injury severity is likely to follow the ordered sequence. Furthermore, the estimation procedure of two ordered models is rather simple while providing more comprehensive information on injury severity than the multinomial one because all levels of injury severity can be evaluated in the ordered model. In the following section, we provide interpretation and discussion on the ordered probit and logit model.

5.2 Marginal Effects and Interpretation

The time from midnight to 6:00 A.M. and the time from 6:00 A.M. to noon were chosen as influential factors. Compared with accidents that occur between noon and midnight, accidents that occur between midnight and 6:00 A.M. are more likely to cause severe injuries and fatalities, whereas accidents that occur between 6:00 A.M. and noon lead to less severe injury levels. As reported in previous research (Eluru *et al*., 2008; Jang *et al*., 2010; Klop, 1998; Kockelman and Kweon, 2002), driving at night leads more severe injuries due to the limited range of visibility, faster vehicle speeds due to less traffic, alcohol use, and driving while drowsy. On the other hand, the traffic between 6:00 A.M. and noon relates mostly to work trips, so drivers are more familiar with their routes, and traffic congestion during the peak hours tends to reduce the possibility of severe crashes.

When an accident occurs on a ramp or at a toll gate, the risk that the accident accompanies with severe injury decreases because most drivers reduce their speed on ramps and at toll gates. Among the primary crash factors, vehicle defects, obstacles

Table 5. Model Estimates of Multinomial Logit

Variables	MNL					
	Fatality			Injury		
	Coefficient	P>|z|	Marginal effect	Coefficient	P>|z|	Marginal effect
Time						
6:00A.M. - Noon	-0.44378	0.027	-5E-05	-0.25992	0.004	-0.01549
Crash location						
Ramp	-0.81	0.03	-7.7E-05	-0.48386	0.001	-0.02592
Toll gate	-2.94402	0	-0.00018	-0.9152	0	-0.04338
Primary crash factor						
Median violation	2.488736	0	0.001015	1.821469	0.002	0.239472
Collision partner						
Car vs. car	1.325507	0	0.00028	0.625398	0	0.048
Car vs. people	6.342891	0	0.022953	3.436344	0	0.604318
No. of parties						
2	1.117615	0	0.000193	0.870885	0	0.069801
3	1.022726	0.005	0.000178	1.159137	0	0.115343
More than 3	1.949684	0	0.000611	1.324806	0	0.142546
Type of vehicle at fault						
Van	0.787953	0.001	0.000135	0.431249	0.001	0.031488

Threshold value	Coefficient	Standard error
ψ_1 (between PDO and Injury)	-2.534318	0.1565934
ψ_2 (between Injury and Fatality)	-3.882193	0.31438

Classification	Number of Independent variables	Log likelihood	Adjusted Likelihood Ratio Index
LL(0)	0	-4,209.4555	-
LL(β)	84	-3,507.9652	0.146691253

*This table shows only the variables that satisfy the 5 % significance level.

and poor road conditions, driving while drowsy, and median violation have been chosen as influential factors. Compared to speeding which is set as a default factor in the category of primary crash factors, factors of vehicle defects, obstacles and poor road conditions tend to decrease the risk of severe injuries. Generally, if an accident is occurred due to one of these factors, it would be easier for the driver to handle the situation and respond so that the accident is avoided. Therefore, it might be reasonable to expect that these factors decrease the severity of injury levels. On the other hand, the marginal effects of driving while drowsy and median violations tend to increase the risk of severe accidents. As expected, drowsy drivers are likely to be unable to make proper response to timely avoid accidents. Even though the frequency of accidents due to median violations is low, such accidents are highly likely to cause fatalities.

Among weather factors, rainy or snowy weather appeared to decrease the injury level. On a rainy or snowy day, drivers usually reduce their speed and drive more carefully than they do on dry days so that the risk of severe injury tends to be diminished.

Since car vs. facility is a default variable in the category of collision partner variables, it is shown to be more probable that accidents result in severe injuries when they occur between car and people (car vs. people), between two cars (car vs. car) and between car and facility (car vs. facility) in order. When there is no collision partner (car only), injury severity tend to be slight compared to car vs. facility collisions.

The number of vehicles involved in an accident is highly relevant to the possibility of a severe injury. As the numbers of cars and occupants of the cars increase, the possibility of the accident producing severe injuries also increases.

Among the types of vehicles at fault, vans are the only type of vehicle that is an influential factor that increases the risk of severe injury levels. Since vans generally are larger and heavier than passenger cars, the impact of collision generally is greater, so the crash partner experiences a greater external force.

The driver's age, geometrical factors (radius of curvature and vertical grade), and arrival time for emergency responders were not influential on the severity of injuries. As reported in previous research (Kockelman and Kweon, 2002; O'Donnell and Connor, 1996; Singleton *et al*., 2004), alcohol use is regarded as one of the major factors that lead to severe injuries. In this study, however, the coefficient of alcohol use variable was not statistically significant in the ordered probit model or in the other two models. This result shows the low tendency of driving under the influence of alcohol on Korean expressways because the expressways are used mainly for long distance travels rather than shorter commute trips.

6. Conclusions

Identifying factors that increase or decrease the risk of severe injuries is one of the fundamental tasks required to enhance the safe operation of Korean expressways. Therefore, this study analyzed traffic crashes that occurred on Korean expressways in 2008 and identified the influential factors that affected the severity of injuries and how they did so. The first step in performing the analysis was to identify and select an appropriate statistical model. In this study, we analyzed crash data by comparatively using ordered probit, ordered logit, and multinomial logit models. Through this procedure, 16 variables were identified and evaluated as major contributing factors to the severity of injuries sustained by drivers involved in crashes on expressways in Korea.

The influential factors that increase the relative risk of a severe accident, compared with the default variable in each corresponding category, were identified as follows: 1) the time between midnight and 6:00 A.M.; 2) driving while drowsy; 3) median violation; 4) car versus car collision; 5) car versus people collision; 6) car only collision; and 7) two or more related vehicles involved; and 8) van. Meanwhile, influential factors that decrease the risk of a severe accident were identified as follows: 1) time between 6:00 A.M. and noon; 2) ramp; 3) toll gate; 4) vehicle defects; 5) obstacles and poor road conditions; and 6) rainy or snowy weather.

By identifying the influential factors on the severity of injury and evaluating how those factors affect the severity of injury, some background information that can be used for policy making and planning for traffic safety can be provided. At the time between midnight and 6:00 A.M., the severity of injury is likely to be greater than at other times of the day. Therefore, effective countermeasures, such as illumination control and installation of warning signs and signals, can be used to lighten the roadway and increase drivers' visibility. For crash location variables, it is common to decelerate the speed of a vehicle inside a ramp or at a toll gate. Thus, some safety countermeasures should be implemented near the locations where abrupt deceleration frequently occurs is needed due to the excessively short radius of curvature of the ramp.

Among primary crash factors, median violation is the most influential factor. If a vehicle crosses the median and causes a head-on collision, the damage is likely to be substantial. One expressway that did not have a barrier in the median had the highest fatality rate due to head-on collisions among all expressways, which indicates that a median barrier likely prevents fatal crashes. Detailed investigation of the causal effect related to this finding is outside the scope of our study. However, this information provides background information that should be used to develop countermeasures to prevent median violation.

Concerning the collision partner variables, the marginal effect of car versus people collision on fatality is fairly high. Since pedestrian access to the expressway is strictly prohibited in Korea except for very limited cases, the car versus people collision does not occur often. Accidents of this type occur when the road constructors and the employees of the expressway corporation are on the road to work or when passengers who already been in an accident step outside their car to wait for help. Although the frequency of this accident type is low, the fatality rate is fairly high due to the characteristics of expressways. Therefore, there is a great need to provide evacuation routes, rescue facilities, and a strategy to keep road constructors and expressway employees safe. The results of this study help transportation planners and policy makers understand which risk factors contribute more to the severity of accidents on the expressways, contribute to better predictions of policy implications, and allow the recommendation and implementation of optimal countermeasures.

Acknowledgements

This work was supported by National Research Foundation of Korea Grant funded by the Korean Government (2009-0075811). The generous support from the Integrated Research Institute of Construction and Environmental Engineering of Seoul National University, and Korea Expressway Corporation is also gratefully acknowledged. The opinions expressed in the paper, however, are solely of the authors and do not necessarily reflect the opinions of the respective agencies.

References

Abdel-Aty, M. (2003). "Analysis of driver injury severity levels at multiple locations using ordered probit models." *Journal of Safety Research,* Vol. 34, No. 5, pp. 597-603.

Abdel-Aty, M. and Abdelwahab, H. (2004). "Modeling rear-end collisions including the role of driver's visibility and light truck vehicles using a nested logit structure" *Accident Analysis & Prevention,* Vol. 36, No. 3, pp. 447-456.

Borooah, V. K. (2001). *Logit and probit: Ordered and multinomial models*. Sage Publications, London.

Eluru, N., Bhat, C. R., and Hensher, D. A. (2008). "A mixed generalized ordered response model for examining pedestrian and bicyclist injury severity level in traffic crashes." *Accident Analysis and Prevention*, Vol. 40, No. 3, pp. 1033-1054.

Jang, K., Park, S. H., Chung, S., and Song, K. (2010). "Influential factors on level of injury in pedestrian crashes: Applications of ordered probit model with robust standard errors." *89th Annual Meeting of the Transportation Research Board*, Washington, D.C.

Klop, J. (1998). "Factors influencing bicycle crash severity on two-lane undivided roadways in North Carolina." *78th Annual Meeting of the Transportation Research Board*, Washington, D.C.

Kockelman, K. M. and Kweon, Y. J. (2002). "Driver injury severity: An application of ordered probit models." *Accident Analysis and Prevention*, Vol. 34, No. 3, pp. 313-321.

Korea Transport Database (2011). *The number of traffic accidents and casualties*, http://www.ktdb.go.kr/.

Lee, H. R., Kim, K. J., and Son S. N. (2011). "A study on the factor analysis by grade for highway traffic accident." *Korean Society of Road Engineers*, Vol. 13, No. 3, pp. 157-165.

Lee, D. M., Kim, E. C., Sung, N. M., and Kim, D. H. (2008). "Development of severity model for rural unsignalized intersecton crashes." *Korean Society of Road Engineers*, Vol. 10, No. 3, pp. 47-56.

Lee, S. I., Lee, H. S., and Won, J. M. (2007). "An analysis of the vehicle characteristics to effect on highway traffic accidents - Focussing on tire accidents." *Korean Society of Civil Engineers*, Vol. 27, No. 4, pp. 405-412.

McKelvey, R. D. and Zavoina, W. (1975). "A statistical model for the analysis of ordinal level dependent variables." *Journal of Mathematical Sociology*, Vol. 4, No. 1, pp. 103-120.

O'Donnell, C. and Connor, D. (1996). "Predicting the severity of motor vehicle accident injuries using models of ordered multiple choices." *Accident Analysis and Prevention*, Vol. 28, No. 6, pp. 739-753.

Oh, J. T. (2006). "Development of severity models for vehicle accident injuries for signalized intersections in rural areas." *KSCE Journal of Civil Engineering*, Vol. 10, No. 3, pp. 219-255.

Oh, J. T., Washington, S., and Lee, D. M. (2010). "Property damage crash equivalency factors to solve crash frequency-severity dilemma: Case study on South Korean rural roads." *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2148, pp. 83-92.

Park, J. T., Kim, Y. S., and Lee, S. B. (2009). "The analysis of older driver's traffic accident characteristic at express-way using Logit model." *Korean Society of Road Engineers*, Vol. 11 No. 4, pp. 1-7.

Park, J. S., Park, G. S., Kim, T. Y., and Park, B. H. (2008). "Characteristics and severity of side right-angle collisions at signalized intersections." *Korean Society of Road Engineers*, Vol. 10, No. 4, pp. 199-211.

Ryu, S. O. and Park, B. H. (2006). "Multivariate regression analysis between geometric structure and traffic accident in expressway intersections." *Journal of the Institute of Constructional Technology*, Vol. 25, No. 1, pp. 1-23.

Shim, K. B. (2009). "The determination of risk group and severity by traffic accidents types - Focusing on Seoul City -." *Korean Society of Road Engineers*, Vol. 11, No. 2, pp. 195-203.

Singleton, M., Qin, H., and Luan, J. (2004). "Factors associated with higher level of injury severity in occupants of motor vehicles that were severely damaged in traffic crashes in Kentucky, 2000-2001." *Traffic Injury Prevention*, Vol. 5, pp. 144-150.

Sohn, S. Y. and Shin, H. W. (2010). "Pattern recognition for road traffic accident severity in Korea." *Ergonomics*, Vol. 44, No. 1, pp. 107-117.

Spiegel, M. R. and Stephens, L. J. (1982). *Schaum's outline of theory and problems of statistics and econometrics*. McGraw-Hill, London.

Statistics Korea (2010). *Annual report on the cause of death statistics (2009)*, Statistics Korea, Korea.

Tarko, A. P., Bar-Gera, H., Thomaz, J., and Issariyanukula, A. (2010). "Model-based application of abbreviated injury scale to police-reported crash injuries." *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2148, pp. 59-68.

Washington, S. P., Karlaftis, M. G., and Mannering, F. L. (2003). *Statistical and econometric methods for transportation data analysis*, Chapman & Hall/CRC, Boca Raton, Florida.