Accident Analysis and Prevention 36 (2004) 869–876

ACCIDENT ANALYSIS & PREVENTION

www.elsevier.com/locate/aap

Bivariate ordered-response probit model of driver's and passenger's injury severities in collisions with fixed objects[☆]

Toshiyuki Yamamoto[a,*], Venkataraman N. Shankar[b,1]

[a] Department of Civil Engineering, Nagoya University, Furo-cho, Chikusa-ku, Nagoya 464-8603, Japan
[b] Department of Civil and Environmental Engineering, University of Washington, 121 More Hall, P.O. Box 352700, Seattle, WA 98195, USA

Received 5 October 2001; received in revised form 12 November 2002; accepted 11 September 2003

Abstract

A bivariate ordered-response probit model of driver's and most severely injured passenger's severity (IS) in collisions with fixed objects is developed in this study. Exact passenger's IS is not necessarily observed, especially when only most severe injury of the accident and driver's injury are recorded in the police reports. To accommodate passenger IS as well, we explicitly develop a partial observability model of passenger IS in multi-occupant vehicle (HOV). The model has consistent coefficients for the driver IS between single-occupant vehicle (SOV) and multiple-occupant vehicle accidents, and provides more efficient coefficient estimates by taking into account the common unobserved factors between driver and passenger IS. The results of the empirical analysis using 4-year statewide accident data in Washington State reveal the effects of driver's characteristics, vehicle attributes, types of objects, and environmental conditions on both driver and passenger IS, and that their IS have different elasticities to some of the risk factors.

Keywords: Bivariate ordered-response probit model; Fixed object; Injury severity; Single-vehicle accident

1. Introduction

The roadside safety problem remains a major problem of highway safety despite dedicated efforts over decades. Each year roadside crashes kill approximately 15,000 people, representing about one third of highway fatalities, and injures another 1,000,000. These crashes cost society an estimated US$ 110 billion. Many of these crashes involve single vehicles that run off the highway and either overturn or crash into a fixed object such as a tree or pole (NCHRP, 2001). Thus, this study, among others, attempts to better understand the nature of single-vehicle accidents in collision with fixed objects.

Many studies have been carried out to examine injury severity (IS) of vehicle accidents including run-off-roadway accidents, in relation to driver factors, vehicle characteristics, and road conditions. Several types of statistical models have been applied to examine accident severity. Many studies have applied methods which take into account the discreteness of the reported IS, where severity is categorised as several predetermined levels. Loglinear models have been applied to examine odds ratios by comparing aggregated counts by some attributes (e.g., Kim et al., 1995; Abdel-Aty et al., 1998). Disaggregate models have also been applied including logistic regressions (Jones and Whitfield, 1988; Lui et al., 1988; Shibata and Fukuda, 1994), ordered-response probit models (O'Donnell and Connor, 1996; Duncan et al., 1998; Khattak, 2001; Kockelman and Kweon, 2002), and multinomial or nested logit models (Shankar et al., 1996, 2000; Chang and Mannering, 1999; Carson and Mannering, 2001; Lee and Mannering, 2002).

The accident severity of interest in these studies mainly concerns the most severe injury in the accident (Shibata and Fukuda, 1994; Shankar et al., 1996; Abdel-Aty et al., 1998; Chang and Mannering, 1999; Carson and Mannering, 2001; Lee and Mannering, 2002), and the driver IS (Lui et al., 1988; Jones and Whitfield, 1988; Khattak, 2001; Kockelman and Kweon, 2002, Shankar et al., 2000). Other than that, IS of front outboard occupant (Farmer et al., 1997) and IS of each occupant (O'Donnell and Connor, 1996) have also been examined in some studies. The most severe injury is our greatest concern, and most data sets on accidents include information on the most severe injury in the accident.

☆ The earlier version of this paper was presented at the 81st Annual Meeting of Transportation Research Board, Washington, DC, 13–17 January 2002.

* Corresponding author. Tel.: +81-52-789-4636; fax: +81-52-789-3738.
E-mail addresses: yamamoto@civil.nagoya-u.ac.jp (T. Yamamoto), vns@u.washington.edu (V.N. Shankar).
[1] Tel: +1-206-616-1259; fax: +1-206-543-1543.

0001-4575/$ – see front matter
doi:10.1016/j.aap.2003.09.002

Vehicle occupancy is an important factor affecting the most severe injury from an accident. It is desirable to develop models for IS by occupant as conducted by O'Donnell and Connor (1996), which can logically derive the most severe injury in the accident. The information on IS by occupant is, however, seldom included in the databases on accidents. Information regarding passengers is particularly inconsistent, especially for passengers who are injured to a lesser extent. Chang and Mannering (1999) developed a joint model of vehicle occupancy and IS of most severely injured occupant, and revealed that some explanatory variables for IS given vehicle occupancy level are significant only at single occupancy. The result suggests that the structures of IS by most severely injured occupant vary according to vehicle occupancy, and that they are different at least between single-occupant vehicle (SOV) and multiple-occupant vehicle (termed HOV in our study). Driver IS can be assumed structurally consistent across vehicle occupancies regardless of the number of passengers, which further substantiates why some of the afore-mentioned studies have used driver IS as a dependent variable.

Hutchinson (1986) developed a bivariate ordered-response probit model of driver and front passenger IS, and examined the effects of overturning with the distinction between rural and urban area on driver and front passenger IS. The parameters were estimated with aggregated counts, assuming only the velocity as an unobserved common factor between the two IS. A bivariate ordered-response probit model of driver's and most severely injured passenger IS is developed in this study, and the parameters are estimated with disaggregate data sets containing many explanatory variables, which enable us to examine the effects of various factors including driver's conditions, vehicle conditions, and road conditions. By controlling for factors relating to the driver, vehicle and roadway, we can be reasonably closer to discerning differences in effects on accident severities due to vehicle occupancy.

First and foremost in this process is the identification of IS. To reiterate, we are interested in modeling jointly driver IS and the most severe passenger IS in a single-vehicle accident. Identification of driver IS is straightforward from existent accident databases. The modeling problem then reduces to assumptions on the most severe passenger IS. Since most severe IS of the accident and driver IS are usually recorded in police reports, exact IS of passengers is not necessarily observed consistently. If the most severe IS of the accident is greater than driver IS, then the most severe IS of passenger is equal to the most severe IS in the accident. Otherwise, IS by most severely injured passenger can be regarded at most the same level of the driver IS. Ordered-response probit models facilitate such partial observability relating to IS level of the most severely injured passenger. While driver IS coefficients can be assumed to be consistent, the partial observability assumption on the limit of passenger IS (i.e. less than or equal to driver IS) allows us use more cases for parameter estimation. The larger sample size, taking into account the common unobserved factors between driver IS and passenger IS, provides more efficient coefficient estimates in the model.

The rest of this paper is organised as follows. The bivariate ordered-response probit model is described in Section 2. This is followed by a description of the data set used in model estimation. The consistency of the univariate ordered-response probit model of most severe IS in the accident between SOV and HOV, and that of driver IS is examined, respectively. Then, the estimation results of the bivariate ordered-response probit model are presented. The last section presents conclusions with a summary.

2. Method

A bivariate ordered-response probit model is an extension of a univariate ordered-response probit model. A univariate ordered-response probit model supposes an unobserved latent and continuous injury severity given as

$$y_i^* = \beta \boldsymbol{x}_i + \varepsilon_i \tag{1}$$

where y_i^* is an unobserved latent and continuous injury severity of accident i, β the vector of parameters to be estimated; and $\boldsymbol{x}_i$ and ε_i are vectors of explanatory variables and a random error term, respectively. IS is observed as a discrete category, y_i, given as

$$y_i = \begin{cases} 0 & \text{if } y_i^* \le 0, \\ 1 & \text{if } 0 < y_i^* \le \mu_1, \\ 2 & \text{if } \mu_1 < y_i^* \le \mu_2, \\ \vdots & \\ J-1 & \text{if } \mu_{J-2} < y_i^* \end{cases} \tag{2}$$

where μ's are unknown parameters to be estimated and J the number of IS categories. In most accident databases, five unique categories exist, namely, property damage only, possible injury, evident injury, disabling injury and fatality. These definitions are consistent with the national fatality analysis reporting system (FARS) structure. Assuming that ε_i is normally distributed, and normalizing the mean and variance of ε_i to 0 and 1 without loss of generality, we have the following probabilities:

$$\begin{aligned}
\Pr(y_i = 0) &= \int_{-\infty}^{-\beta \boldsymbol{x}_i} \phi(\varepsilon_i)\,\mathrm{d}\varepsilon_i = \Phi(-\beta \boldsymbol{x}_i),\\
\Pr(y_i = 1) &= \int_{-\beta \boldsymbol{x}_i}^{\mu_1-\beta \boldsymbol{x}_i} \phi(\varepsilon_i)\,\mathrm{d}\varepsilon_i = \Phi(\mu_1-\beta \boldsymbol{x}_i) - \Phi(-\beta \boldsymbol{x}_i),\\
\Pr(y_i = 2) &= \int_{\mu_1-\beta \boldsymbol{x}_i}^{\mu_2-\beta \boldsymbol{x}_i} \phi(\varepsilon_i)\,\mathrm{d}\varepsilon_i\\
&= \Phi(\mu_2-\beta \boldsymbol{x}_i) - \Phi(\mu_1-\beta \boldsymbol{x}_i),\\
&\vdots\\
\Pr(y_i = J-1) &= \int_{\mu_{J-2}-\beta \boldsymbol{x}_i}^{\infty} \phi(\varepsilon_i)\,\mathrm{d}\varepsilon_i = 1 - \Phi(\mu_{J-2}-\beta \boldsymbol{x}_i)
\end{aligned} \tag{3}$$

where ϕ and Φ are the standard normal probability density and cumulative distribution functions, respectively. Even if IS is observed only partially (for example, $y_i \le 1$, or $1 \leq y_i \leq 2$, etc.) such cases are easily treated in an ordered-response probit model, and the generalized form of the probability of observing a case is given as

$$\Pr(k_i \leq y_i \leq l_i) = \int_{\mu_{k_i-1}-\beta\boldsymbol{x}_i}^{\mu_{l_i}-\beta\boldsymbol{x}_i} \phi(\varepsilon_i)\,\mathrm{d}\varepsilon_i = \Phi(\mu_{l_i} - \beta\boldsymbol{x}_i) - \Phi(\mu_{k_i-1} - \beta\boldsymbol{x}_i) \qquad (4)$$

where μ_{-1}, μ_0, and μ_{J-1} are $-\infty$, 0, and ∞, respectively. The case for $y_i \leq l_i$ is represented mathematically by the same notation, $0 \leq y_i \leq l_i$ in the equation.

A bivariate ordered-response probit model considers two dependent variables, y_{1i} and y_{2i}, and assumes that the two random error terms, ε_{1i} and ε_{2i} form a bivariate normal distribution with correlation between them. The probability is given as:

$$\begin{aligned}
&\Pr(k_i \leq y_{1i} \leq l_i, m_i \leq y_{2i} \leq n_i) \\
&= \int_{\mu_{2,m_i-1}-\beta_2\boldsymbol{x}_{2i}}^{\mu_{2n_i}-\beta_2\boldsymbol{x}_{2i}} \int_{\mu_{1,k_i-1}-\beta_1\boldsymbol{x}_{1i}}^{\mu_{1l_i}-\beta_1\boldsymbol{x}_{1i}} \phi_2(\varepsilon_{1i}, \varepsilon_{2i}, \rho)\,\mathrm{d}\varepsilon_{1i}\,\mathrm{d}\varepsilon_{2i} \\
&= \Phi_2(\mu_{1l_i} - \beta_1\boldsymbol{x}_{1i}, \mu_{2n_i} - \beta_2\boldsymbol{x}_{2i}, \rho) \\
&\quad -\Phi_2(\mu_{1,k_i-1} - \beta_1\boldsymbol{x}_{1i}, \mu_{2n_i} - \beta_2\boldsymbol{x}_{2i}, \rho) \\
&\quad -\Phi_2(\mu_{1l_i} - \beta_1\boldsymbol{x}_{1i}, \mu_{2,m_i-1} - \beta_2\boldsymbol{x}_{2i}, \rho) \\
&\quad + \Phi_2(\mu_{1,k_i-1} - \beta_1\boldsymbol{x}_{1i}, \mu_{2,m_i-1} - \beta_2\boldsymbol{x}_{2i}, \rho) \qquad (5)
\end{aligned}$$

where ϕ_2 and Φ_2 are the standard bivariate normal probability density and cumulative distribution functions; ρ is an unknown correlation between ε_{1i} and ε_{2i} to be estimated; and $\mu_{a,-1}$, $\mu_{a,0}$, and $\mu_{a,Ja-1}$ are $-\infty$, 0, and ∞, respectively, for $a = 1, 2$. Supposing y_{1i} and y_{2i} are the discrete observed categories of IS of driver and most severely injured passenger, respectively, log-likelihood of joint model considering both SOV and HOV accidents is given as

$$\begin{aligned}
\mathrm{LL} = &\sum_{i\in\mathrm{SOV}} \ln\{\Phi(\mu_{1l_i} - \beta_1\boldsymbol{x}_{1i}) - \Phi(\mu_{1,k_i-1} - \beta_1\boldsymbol{x}_{1i})\} \\
&+ \sum_{i\in\mathrm{HOV}} \ln\{\Phi_2(\mu_{1l_i} - \beta_1\boldsymbol{x}_{1i}, \mu_{2n_i} - \beta_2\boldsymbol{x}_{2i}, \rho) \\
&-\Phi_2(\mu_{1,k_i-1} - \beta_1\boldsymbol{x}_{1i}, \mu_{2n_i} - \beta_2\boldsymbol{x}_{2i}, \rho) \\
&-\Phi_2(\mu_{1l_i} - \beta_1\boldsymbol{x}_{1i}, \mu_{2,m_i-1} - \beta_2\boldsymbol{x}_{2i}, \rho) \\
&+\Phi_2(\mu_{1,k_i-1} - \beta_1\boldsymbol{x}_{1i}, \mu_{2,m_i-1} - \beta_2\boldsymbol{x}_{2i}, \rho)\} \qquad (6)
\end{aligned}$$

where SOV and HOV are the single-occupant vehicle and multi-occupant vehicle accidents, respectively. The unknown parameters, β_1, β_2, $\mu_{a,j}$ ($j = 1, 2, \ldots, J_a - 2$, $a = 1, 2$), and ρ, are simultaneously estimated by the maximum likelihood method. As mentioned in Section 1, the driver IS coefficients are consistent across occupancy (SOV and HOV), thus the same parameter vector, β_1, is used for the driver IS equations of the single-occupant vehicle and multi-occupant vehicle accidents. Once the parameters are estimated, driver IS of the single-occupant vehicle accident is predicted by the first term, and driver and passenger IS of the multi-occupant vehicle accident is predicted by the latter term. The positive (negative) value of the coefficient estimate means that the probability of high IS increases (decreases) with the variable value.

3. Empirical data set

The source of the data used for the model estimation in this study is the Washington State accident records database, which is supplied by the Washington State Department of Transportation. From the database containing many types of accident, data on single-vehicle collisions with fixed objects were used to estimate the model. The records contain a variety of information including weather conditions, roadway conditions, object type, driver's conditions, vehicle conditions as well as IS reported at the time of the accident. The explanatory variables used in the empirical analysis of this study are summarized in Table 1. Unfortunately, the data set doesn't include passenger conditions, such as passenger location, gender and age. Almost all the sub samples divided by the dummy variables has sufficient cases. All the sub sample sizes are more than 50 cases, except one divided by over centerline indicator with 19 cases.

The most severe IS in the accident and the driver IS are recorded, and the number of persons injured in the accident and that of fatalities are also recorded. Information on passenger IS for every passenger is not consistently available, thus leading to partial observability on passenger IS. IS is numerically categorized as (0) property damage only, (1) possible injury, (2) evident injury, (3) disabling injury, and (4) fatality. A total of 20,363 accidents reported during the 4-year period from 1993 to 1996 were used in this study, with 11,505 (56.5%) accidents being property damage only. Other most severe IS distributions included 3373 (16.6%) possible injury, 4318 (21.2%) evident injury, 965 (4.7%) disabling injury, and 202 (1.0%) fatal injuries, respectively.

The most severe IS in the accident is greater than driver IS in about 20% (1376 out of 6647) of all multiple-occupant vehicle accidents, so the passenger IS is equal to the most severe IS in the accident in these cases. The passenger IS for the rest of the sample cases is not directly obtainable and only regarded as not greater than driver IS. To obtain IS distributions of the most severely injured passengers for these cases, the lower and upper bounds of possible IS by most severely injured passenger are derived. The information on the number of injured persons in the accident and that of fatalities are incorporated with the censored observation on the passenger IS to obtain the bounds. We assume that the most severe IS, the number of injured persons, and the number of fatalities in the accidents exclude pedestrian injuries, and include only occupants' injury. The validity of the assumption is indirectly examined by the rate of pedestrian injuries in SOV accidents, where reported injury to non drivers accounts for only 0.9% (=126/13,716) of all SOV accidents. The result suggests that the possible bias resulting from the assumption is negligible.

Table 1
Description of explanatory variable

Variable	Description	Mean
Road condition		
Year 95 indicator	1 if accident occurred in 1995, 0 otherwise	
Summer indicator	1 if accident occurred in the summer, 0 otherwise	0.2611
Spring or fall indicator	1 if accident occurred in the spring or fall, 0 otherwise	0.2959
Principal arterial state highway indicator	1 if accident occurred on the principal arterial state highway, 0 otherwise	0.3941
Minor collector state highway indicator	1 if accident occurred on the minor collector state highway, 0 otherwise	0.1388
Collector state highway indicator	1 if accident occurred on the collector state highway, 0 otherwise	0.0870
Intersection indicator	1 if accident occurred at intersection and related, 0 otherwise	0.0693
Intersection-related indicator	1 if accident occurred at intersection-related but not at intersection, 0 otherwise	0.0068
Straight and grade roadway indicator	1 if roadway is straight and grade, 0 otherwise	0.1879
Straight and sag indicator	1 if roadway is straight and sag, 0 otherwise	0.0075
Curve and level roadway indicator	1 if roadway is curve and level, 0 otherwise	0.1511
Curve and grade roadway indicator	1 if roadway is curve and grade, 0 otherwise	0.2119
Curve and hill crest indicator	1 if roadway is curve and hill crest, 0 otherwise	0.0047
Bridge indicator	1 if roadway is located on bridge, overpass or ferry dock, 0 otherwise	0.0490
Underpass indicator	1 if accident occurred at underpass or tunnel, 0 otherwise	0.0071
Off roadway indicator	1 if accident occurred off roadway, 0 otherwise	0.9656
Wet roadway surface indicator	1 if roadway surface condition was wet, 0 otherwise	0.2708
Icy roadway surface indicator	1 if roadway surface condition was ice or snow, 0 otherwise	0.2355
Rain indicator	1 if the weather was raining, 0 otherwise	0.2068
Snow indicator	1 if the weather was snow, 0 otherwise	0.0617
Dark indicator	1 if accident occurred in dirk time with no street lights, 0 otherwise	0.2406
Posted speed	in 100 mile/h	0.5246
Fixed object		
Post indicator	1 if accident occurred in collision with wood and metal sign post, guide post, 0 otherwise	0.0427
Ditch indicator	1 if accident occurred in collision with culvert end or other appurtenance in ditch, roadway ditch, 0 otherwise	0.0797
Guardrail end indicator	1 if accident occurred in collision with leading end of guardrail, 0 otherwise	0.0113
Guardrail face indicator	1 if accident occurred in collision with face of guardrail, 0 otherwise	0.1564
Concrete barrier face indicator	1 if accident occurred in collision with face of concrete barrier, 0 otherwise	0.1827
Bridge face indicator	1 if accident occurred in collision with face of bridge, 0 otherwise	0.0441
Construction machinery indicator	1 if accident occurred in collision with road or construction machinery, 0 otherwise	0.0087
Tree indicator	1 if accident occurred in collision with tree or stump, 0 otherwise	0.0529
Fence indicator	1 if accident occurred in collision with fence, 0 otherwise	0.0356
Mail box indicator	1 if accident occurred in collision with mail box, 0 otherwise	0.0086
Vehicle condition		
Lap belt indicator	1 if lap belt was used, 0 otherwise	0.0694
Shoulder belt indicator	1 if shoulder belt was used, 0 otherwise	0.0079
Lap and shoulder belt indicator	1 if lap and shoulder belt are used, 0 otherwise	0.7798
Air bag and belt indicator	1 if air bag and belt are used, 0 otherwise	0.0147
Vehicle age	in 10 years	0.9219
Large truck indicator	1 if vehicle was truck over 10 K pounds, 0 otherwise	0.0026
Truck indicator	1 if vehicle was truck tractor, semi trailer or other truck combinations, 0 otherwise	0.0331
Motorcycle indicator	1 if vehicle is motorcycle, scooter bike or moped, 0 otherwise	0.0044
Defective brake indicator	1 if brakes were defective, 0 otherwise	0.0110
Tire blow indicator	1 if tires had punctured or blown, 0 otherwise	0.0140
Worn tire indicator	1 if tire was worn or smooth, 0 otherwise	0.0247
Lost wheel indicator	1 if wheel was lost, 0 otherwise	0.0039
Defect indicator	1 if steering mechanism was defective, 0 otherwise	0.0065
Driver condition		
Male driver indicator	1 if driver was male, 0 otherwise	0.6608
Driver's age	in 100 years	0.3427
Elderly driver indicator	1 if driver's age was 55 and over, 0 otherwise	0.1112
Over speed limit indicator	1 if speed exceeded stated speed limit, 0 otherwise	0.0277
Exceed safety speed indicator	1 if speed exceeded reasonable safe speed, 0 otherwise	0.4364
Inattention indicator	1 if driver was in inattention, 0 otherwise	0.1130
Defective equipment indicator	1 if driver was operating defective equipment, 0 otherwise	0.0386
Improper turning indicator	1 if driver made improper turning, 0 otherwise	0.0075
Over centerline indicator	1 if driver drove over centerline, 0 otherwise	0.0009
Asleep indicator	1 if driver was apparently asleep, 0 otherwise	0.0908
Driver sobriety indicator	1 if driver had been drinking and ability was impaired, 0 otherwise	0.1518
Number of passengers		0.5076

Table 2
Accident cross-classification frequencies by driver's severity and most severely injured passenger in multiple-occupant accidents

		Most severely injured passenger's severity[a]							
		0	1	1–2[b]	2	1–3[b]	3	4	Total
	0	3271	528	–	447	–	47	1	4294
Driver's	1	278	435	–	170	–	35	7	925
severity[a]	2	258	–	751	–	–	109	14	1132
	3	29	–	–	–	212	–	18	259
	4	1	–	–	–	23	–	13	37
	Total	3837	963	751	617	235	432	53	6647

[a] Severity is categorized as 0: not injured; 1: possible injury; 2: evident injury; 3: disabling injury; and 4: fatality.
[b] IS is reported as censored, and derived as the lower and upper bound of possible IS.

The cross table of driver IS and the most severely injured passenger IS in multiple-occupant vehicle accidents is shown in Table 2. The table shows that more passengers than drivers are killed (53 > 37), and less passengers than drivers are not injured (3837 < 4294) in the HOV accidents. These can partly be explained by the presence of multiple passengers, which multiply the likelihood of injury of passengers. The average number of passengers in HOV accidents is 1.55, and 65.9% of such accidents had one passenger in the sample.

4. Model estimation

Before estimating the bivariate ordered-response probit model proposed in this study, the univariate ordered-response probit models of most severe IS and driver IS are estimated to examine the consistency of the models across occupancy (SOV versus HOV), as well as geographic effects (urban versus rural). A statistically significant difference between urban and rural areas was confirmed, and separate models according to the distinction were developed in previous studies (e.g. Lee and Mannering, 2002). The likelihood ratio test (Greene, 2000) is carried out by comparing the values of the log-likelihood function at convergence with and without the restrictions imposed. The test statistic is given as

$$\mathrm{LR} = -2\{\mathrm{LL_r}(\beta) - \mathrm{LL_{s1}}(\beta) - \mathrm{LL_{s2}}(\beta)\} \tag{7}$$

where $\mathrm{LL_r}(\beta)$ is the log-likelihood at convergence of the model estimated on the pooled sample; and $\mathrm{LL_{s1}}(\beta)$ and $\mathrm{LL_{s2}}(\beta)$ are the log-likelihood at convergence of the models estimated on separate samples. The statistic is χ^2 distributed with the degrees of freedom equal to the sum of the number of parameters estimated in the models on separate samples minus that on pooled sample.

Statistical difference is tested for in two different severity cases, most severe IS of accident, and driver IS. Within these IS, occupancy effects (SOV versus HOV) and spatial effects (urban versus rural) are examined. The results of these tests are shown in Table 3. The results suggest for either IS dimension (most severe IS in accident or driver IS), that the

Table 3
Consistency tests for occupancy and spatial effect differences in explanatory effects on most severe injury and driver injury

	$\mathrm{LL_r}(\beta)$[a]	$\mathrm{LL}(\beta)$		χ^2 (d.f.)
SOV/HOV		SOV	HOV	
Most severe injury	−21300	−13650	−7593	114 (d.f. = 47[b])**
Driver injury	−19576	−13408	−6131	74 (d.f. = 48)**
Urban/rural		Urban	Rural	
Most severe injury	−21300	−9986	−11232	164 (d.f. = 48[b])**
Driver injury	−19576	−9194	−10300	164 (d.f. = 49)**

[a] Log-likelihood at convergence of the model estimated on the pooled sample.
[b] The number of passengers variable is included in the pooled model, but excluded from the single-occupant model.
** $P < 0.01$.

hypothesis that explanatory effects (such as driver, vehicle and roadway factors) are consistent between single-occupant and multi-occupant accidents can be rejected at the 99% confidence level. Likewise, explanatory effects are not consistent between urban and rural areas at the 99% confidence level. The likelihood-ratio-values also suggest that the spatial differences are stronger compared to occupancy effects.

We now examine whether occupancy effects nested within a particular spatial dimension are significant. For example, are explanatory effects for single-occupant collisions with fixed objects in urban areas statistically consistent with those for multi-occupant collisions with fixed objects in urban areas? We examine this hypothesis for both the most severe IS of the accident and driver IS. The results of these tests are given in Table 4. The finding is that driver, vehicle and roadway effects are consistent for driver IS, while inconsistent for most severe IS in the accident. Likewise, similar findings are obtained for collisions with fixed objects in rural areas. The results are consistent with the finding by Chang and Mannering (1999) that most severe IS effects are not transferable across single and multi-occupant scenarios, and confirm the validity of the assumption in this study on the consistency of the model of driver IS.

Table 4
Consistency tests for occupancy effects nested within urban and rural spatial dimensions

	$\mathrm{LL_r}(\beta)$[a]	$\mathrm{LL}(\beta)$		χ^2 (d.f.)
		SOV	HOV	
Urban				
Most severe injury	−9998	−6915	−3058	50 (d.f. = 32)*
Driver injury	−9199	−6748	−2429	44 (d.f. = 34)
Rural				
Most severe injury	−11234	−6679	−4508	94 (d.f. = 40)**
Driver injury	−10292	−6593	−3671	56 (d.f. = 46)

[a] Log-likelihood at convergence of the model estimated on the pooled data.
* $P < 0.05$.
** $P < 0.01$.

Table 5
Summary statistics of the model estimation for urban and rural area

	Urban area	Rural area
Sample size	9723	10640
Parameter estimate of ρ	0.543	0.623
t-Statistic for $\rho=0$	25.20	39.85
LL (0)	−19737	−22956
LL (β_r)[a]	−11633	−13969
LL (β)	−11507	−13781
χ^2	16460 (d.f. = 61)	18349 (d.f. = 77)
$-2\{LL(\beta_r) - LL(\beta)\}$	252 (d.f. = 1)	376 (d.f. = 1)

[a] The log-likelihood at convergence obtained by setting $\rho = 0$.

To further investigate severity models across urban and rural dimensions, we estimate bivariate ordered-response probit models of driver IS and the most severely injured passenger IS. As noted in an earlier section, the likelihood of correlation due to common unobserved factors between driver and passenger IS is significant. Estimating models without accounting for such correlation would result in inefficient parameter estimates. To illustrate this, separate two univariate ordered-response probit models (without the correlation parameter) for driver IS and most severe passenger IS are also estimated. Summary statistics of the models are presented in Table 5. First and foremost, the table shows that estimates of ρ, the correlation parameter, are statistically

Table 6
Coefficient estimates for urban model

Variable	Driver IS			Passenger IS		
	Coefficient	t-Statistic	E[a]	Coefficient	t-Statistic	E[a]
Summer indicator				0.127	2.67	0.44
Intersection indicator	−0.145	−3.25	−0.39	−0.283	−3.55	−0.68
Intersection-related indicator	−0.454	−3.03	−0.80			
Straight and grade roadway indicator				−0.108	−1.99	−0.31
Straight and sag indicator				0.471	2.24	3.03
Bridge indicator				0.499	1.98	3.23
Off roadway indicator	0.581	7.45	1.04			
Icy roadway surface indicator	−0.195	−4.85	−0.52			
Rain indicator	−0.137	−4.55	−0.40			
Post indicator	−0.458	−6.82	−0.88	−0.532	−3.79	−0.97
Ditch indicator	−0.276	−4.09	−0.63			
Guardrail end indicator	0.216	2.13	0.93			
Guardrail face indicator	−0.169	−4.54	−0.46			
Concrete barrier face indicator	−0.136	−4.39	−0.41			
Bridge face indicator	−0.151	−2.57	−0.39	−0.431	−1.60	−0.85
Construction machinery indicator				−0.583	−1.84	−0.91
Tree indicator	0.174	2.76	0.70	0.426	3.32	2.53
Fence indicator	−0.304	−3.71	−0.65	−0.478	−2.63	−0.86
Lap belt indicator	−0.791	−12.69	−1.42	−0.579	−5.04	−1.06
Shoulder belt indicator	−0.552	−4.32	−0.86			
Lap and shoulder belt indicator	−1.007	−27.18	−13.38	−0.680	−11.22	−5.38
Air bag and belt indicator	−0.379	−4.02	−0.73			
Vehicle age	−0.066	−3.41	−0.19			
Large truck indicator	−0.637	−2.27	−0.91			
Truck indicator	−0.188	−2.20	−0.46	−0.762	−2.18	−1.01
Motorcycle indicator	−0.373	−1.41	−0.71			
Defective brake indicator	−0.246	−2.25	−0.56			
Tire blow indicator	−0.244	−2.42	−0.56	−0.482	−2.88	−0.86
Lost wheel indicator				−0.805	−2.89	−1.02
Defect indicator				−0.540	−2.37	−0.90
Male driver indicator	−0.27	−10.2	−1.04	0.095	2.07	0.30
Driver's age	0.181	2.06	0.19			
Elderly driver indicator				0.249	3.13	1.10
Over speed limit indicator	0.145	2.10	0.55			
Exceed safety speed indicator	0.09	2.92	0.29			
Improper turning indicator	−0.445	−2.72	−0.80			
Asleep indicator	0.291	5.29	1.37	0.431	3.72	2.57
Driver sobriety indicator	0.453	12.6	2.60	0.353	5.49	1.77
Number of passengers	−0.037	−2.55	−0.05	0.172	7.32	0.81
Constant	0.219	2.38	–	−0.0104	−0.14	–
μ_1	0.567	43.6	–	0.674	22.0	–
μ_2	1.652	64.0	–	1.776	30.3	–
μ_3	2.486	52.1	–	2.623	21.5	–

[a] The sample means of elasticity and pseudo-elasticity of fatality.

highly significant for both urban and rural area, and that the error terms of driver IS and the most severely injured passenger IS are positively correlated with each other. The table also shows that the log-likelihood at convergence improves statistically significantly for both urban (−11,633 versus −11,507) and rural area (−13,969 versus −13,781) by accounting for correlation between driver and most severe passenger IS. It should also be noted (results not presented) that most of the parameter estimates have smaller standard errors in the correlational model compared to the restricted model ($\rho = 0$). These results suggest that the proposed model is more efficient than the restricted model.

For the purpose of investigative insight into the nature of explanatory effects, we present detailed results of the urban correlational models in Table 6. The detailed results of the rural correlational models are omitted, but can be obtained from the web site (http://www.trans.civil.nagoya-u.ac.jp/~toshi/papers/Tables.pdf). To summarize, the similar results are obtained from both the urban and rural correlational models. For the first attempt, only main effects of the explanatory variables are included in the present model, and the interaction effects or non-linear effects of these variables should be examined in the further study. Examining the coefficient estimates for driver IS in the urban model, the intersection-related indicators have negative coefficient estimates, suggesting higher probability of lower driver IS at intersection locations. These may imply lower speed when approaching intersections and more attention by drivers. The coefficient estimate of the off-roadway indicator indicates higher probability of severe driver IS. Icy roadway surfaces and rain decrease the probability of severer driver IS, because they tend to lower vehicle speeds. Type of fixed objects affects driver IS significantly as well; collisions with leading ends of guardrail and trees tend to cause severe driver IS, while collisions with sign posts, appurtenances in ditch, faces of guardrail, concrete barrier or bridge, and fences tend to cause less severe driver IS. Proper use of restraint system highly significantly decreases the probability of severe driver IS. The results suggest that the restraint with both lap and shoulder belt has a larger effect than that with lap belt alone or with shoulder belt alone, which seems reasonable. On the other hand, the results suggest that the restraint with shoulder belt has a larger effect than that with lap belt, and that the restraint with both air bag and belt has a smaller effect than others. These results seem unreasonable, and needs further investigations on the effects of the restraint systems (the same results on the effects of the restraint systems are also shown for the rural correlational model). The results should be also compared with published effectiveness measures for restraint systems then. Male and younger drivers have lower probability of severe IS. Higher vehicle speed, driver being asleep, or under the influence of alcohol cause more severe driver IS as expected. Vehicle defect variables are also significant.

Significant effects in the most severely injured passenger IS equation are similar to that of driver IS, although fewer attributes are included for explanatory variables. Intersection indicator, the type of fixed objects, proper use of restraint system, driver being asleep, or under the influence of alcohol all have the same effects on the most severely injured passenger IS as they do on driver injury. The effects of the restraint with shoulder belt and that with both air bag and belt, which turned out to be unreasonable in the driver IS equation, are not included in the passenger IS equation, because of the insignificant coefficient estimates. Interestingly, the coefficient for male drivers has a positive coefficient estimate, an opposite sign from driver IS. As mentioned in the former section, passenger conditions are not included in the data set, thus driver conditions may play as a proxy of passenger conditions in the passenger IS equation. If this is the case, a male-female pair as the driver-passenger pair in HOV might explain the results. The coding bias in female is another possible explanation. However, these explanations need a further research with data sets including passenger conditions. Elderly driver indicator has a positive coefficient estimate as well. It may imply that the driver-passenger pair tends to be included in the same age category, and that elderly passenger tends to have severer IS. As in the driver IS equation, vehicle defect variables are significant as well. The number of passenger has a positive coefficient estimate in the passenger equation, suggesting that the probability of higher IS by the most severely injured passenger increases along the number of passengers. The result is reasonable because each passenger has a probability of high IS, and the highest IS of the passengers are treated as most severely injured passenger IS in the model.

To provide some insight into the implications of the estimation results, elasticities of IS probabilities are computed for both driver IS and the most severely injured passenger IS. The elasticity is given as

$$E_{x_{ain}}^{\Pr(y_{ai}=j)} = \frac{\partial \ln \Pr(y_{ai}=j)}{\partial \ln x_{ain}} = \frac{\phi(\mu_{a,j-1} - \beta_a \boldsymbol{x}_{ai}) - \phi(\mu_{aj} - \beta_a \boldsymbol{x}_{ai})}{\Phi(\mu_{aj} - \beta_a \boldsymbol{x}_{ai}) - \Phi(\mu_{a,j-1} - \beta_a \boldsymbol{x}_{ai})} \beta_{an} x_{ain} \tag{8}$$

where x_{ain} and β_{an} are the nth explanatory variable of accident i and the corresponding coefficient estimate of driver IS ($a = 1$) and the most severely injured passenger IS ($a = 2$). The elasticity is, however, appropriate only for continuous variables, and a pseudo-elasticity is calculated for the indicator variables to provide an approximate elasticity. The pseudo-elasticity is defined in this study as

$$E_{x_{ain}}^{\Pr(y_{ai}=j)} = \frac{\begin{array}{l}\{(\Phi[\mu_{aj} - \{\beta_a \boldsymbol{x}_{ai} + \beta_{an}(1 - x_{ain})\}] - \Phi[\mu_{a,j-1} - \{\beta_a \boldsymbol{x}_{ai} + \beta_{an}(1 - x_{ain})\}]) \\ -[\Phi\{\mu_{aj} - (\beta_a \boldsymbol{x}_{ai} - \beta_{an} x_{ain})\} - \Phi\{\mu_{a,j-1} - (\beta_a \boldsymbol{x}_{ai} - \beta_{an} x_{ain})\}]\}\end{array}}{\Phi(\mu_{aj} - \beta_a \boldsymbol{x}_{ai}) - \Phi(\mu_{a,j-1} - \beta_a \boldsymbol{x}_{ai})} \tag{9}$$

The elasticities and pseudo-elasticities are computed for each category of IS and for each variable, and only those of fatality are examined here in particular. The sample means of elasticities and pseudo-elasticities of fatality for urban model are shown in Table 6. Full matrices of the elasticities and pseudo-elasticities can be obtained from the web site (http://www.trans.civil.nagoya-u.ac.jp/∼toshi/papers/Tables.pdf). Both driver IS and the most severely injured passenger IS have high elasticity to the proper use of lap and shoulder belt variable, suggesting the probability of fatality decreases by significant amounts (elasticity of −13.38) if restraint system is used properly. Among the types of fixed objects, tree indicator has the largest elasticity for the most severely injured passenger IS. The result suggests that the collision with trees increases the probability of passenger's fatality by about 250%.

5. Summary and conclusion

Many studies have been carried out to examine IS of vehicle accidents in relation to driver factors, vehicle attributes, and roadway conditions. These studies have relied on the main measures of accident IS, the most IS in the accident, and driver IS. Most severe passenger IS is not known consistently. Little is known on the simultaneous impacts of driver, vehicle and roadway factors on driver IS and most severe passenger IS. This study attempts to investigate that simultaneity by estimating a bivariate ordered-response probit model of driver IS and the most severely injured passenger IS. The ordered-response framework allows us to account for partial observability in passenger IS, while improving efficiency of parameter estimates.

In estimating the bivariate correlational models of driver IS and most severe passenger IS, we note significant shifts in IS distributions along the dimensions of vehicle occupancy and space (urban area versus rural area). Detailed examination of the bivariate model of urban area fixed object collisions indicate that the error terms of driver IS and the most severely injured passenger IS are highly and positively correlated with each other, and that the proposed model is more efficient than conventional univariate models. The estimation results also reveal the effects of driver characteristics, vehicle attributes, types of fixed objects, and environmental conditions on both driver and passenger IS. Further, elasticity computations of key variables provide insight into their marginal impact and hence useful design policy direction. The models estimated in this paper are applicable to other types of accidents. Examining the efficiency of bivariate models is still needed in a variety of contexts, along with a thorough investigation of interactions of key factors. In combination, these issues will enhance our understanding of simultaneous accident structures. Specifically, different structures in the most severely injured passenger IS model, as a function of the number of passengers may increase the preciseness of the models, but may also require larger sample sizes.

Acknowledgements

This study was conducted during the first author's stay at the University of Washington, Seattle as a visiting scholar hosted by Fred L. Mannering, and supported, in part, by a research grant to the first author by the Konoike Foundation. The authors acknowledge anonymous reviewers for their valuable comments.

References

Abdel-Aty, M.A., Chen, C., Schott, J.R., 1998. An assessment of the effect of driver age on traffic accident involvement using log-linear models. Acc. Anal. Prev. 30, 851–861.

Carson, J., Mannering, F., 2001. The effect of ice warning signs on accident frequencies and severities. Acc. Anal. Prev. 33, 99–109.

Chang, L.-Y., Mannering, F., 1999. Analysis of injury severity and vehicle occupancy in truck- and non-truck-involved accident. Acc. Anal. Prev. 31, 579–592.

Duncan, C., A. Khattak, F. Council, 1998. Applying the ordered probit model to injury severity in truck-passenger car rear-end collisions. Transport. Res. Rec. 1635, 63–71.

Farmer, C.M., Braver, E.R., Mitter, E.L., 1997. Two-vehicle side impact crashes: the relationship of vehicle and crash characteristics to injury severity. Acc. Anal. Prev. 29, 399–406.

Greene, W., 2000. *Econometric Analysis*, 4th edition, Prentice-Hall, NJ.

Hutchinson, T.P., 1986. Statistical modelling of injury severity, with special reference to driver and front seat passenger in single-vehicle crashes. Acc. Anal. Prev. 18, 157–167.

Jones, I.S., Whitfield, R.A., 1988. Predicting injury risk and risk-taking behaviour among young drivers. Acc. Anal. Prev. 18, 411–419.

Khattak, A.J., 2001. Injury Severity in Multi-Vehicle Rear-End Crashes. Transport. Res. Rec. 1746, 59–68.

Kim, K., Nitz, L., Richardson, J., Li, L., 1995. Personal and behavioral predictors of automobile crash and injury severity. Acc. Anal. Prev. 27, 469–481.

Kockelman, K.M., Kweon, Y.-J., 2002. Driver injury severity: an application of ordered probit models. Acc. Anal. Prev. 34, 313–321.

Lee, J., Mannering, F., 2002. Impact of roadside features on the frequency and severity of run-off-roadway accidents: an empirical analysis. Acc. Anal. Prev. 34, 149–161.

Lui, K.-J., McGee, D., Rhodes, P., Pollock, D., 1988. An application of a conditional logistic regression to study the effects of safety belts, principal impact points, and car weights on drivers' fatalities. J. Saf. Res. 19, 197–203.

NCHRP, 2001. Strategic plan for improving roadside safety. NCHRP Web Document 33 (Project G17-13): Contractor's Final Report. Transportation Research Board, National Research Council, Washington, DC.

O'Donnell, C.J., Connor, D.H., 1996. Predicting the severity of motor vehicle accident injuries using models of ordered multiple choice. Acc. Anal. Prev. 28, 739–753.

Shankar, V., Mannering, F., Barfield, W., 1996. Statistical analysis of accident severity on rural freeways. Acc. Anal. Prev. 28, 391–401.

Shankar, V., Albin, R., Milton, J., Nebergall, M., 2000. In-Service Performance-Based Roadside Design Policy: Preliminary Insights from Washington State's Bridge Rail Study. Transport. Res. Rec. 1720, 72–79.

Shibata, A., Fukuda, K., 1994. Risk factors of fatality in motor vehicle traffic accidents. Acc. Anal. Prev. 26, 391–397.