Accident Analysis and Prevention 66 (2014) 120–135

Contents lists available at ScienceDirect

Accident Analysis and Prevention

journal homepage: www.elsevier.com/locate/aap

Examining driver injury severity in two vehicle crashes – A copula based approach

Shamsunnahar Yasmin[a,1], Naveen Eluru[a,*], Abdul R. Pinjari[b,2], Richard Tay[c,3]

[a] Department of Civil Engineering & Applied Mechanics, McGill University, Suite 483, 817 Sherbrooke St. W., Montréal, Canada
[b] Department of Civil and Environmental Engineering, University of South Florida, 4202 E. Fowler Avenue, Tampa, FL 33620, United States
[c] Faculty of Business, Economics and Law, La Trobe University, Melbourne, Victoria 3086, Australia

ARTICLE INFO

Article history:
Received 5 August 2013
Received in revised form 6 January 2014
Accepted 20 January 2014

Keywords:
Injury severity profiles
Copula methods
Joint multinomial logit and ordered logit model
Two vehicle crashes

ABSTRACT

A most commonly identified exogenous factor that significantly affects traffic crash injury severity sustained is the collision type variable. Most studies consider collision type only as an explanatory variable in modeling injury. However, it is possible that each collision type has a fundamentally distinct effect on injury severity sustained in the crash. In this paper, we examine the hypothesis that collision type fundamentally alters the injury severity pattern under consideration. Toward this end, we propose a joint modeling framework to study collision type and injury severity sustained as two dimensions of the severity process. We employ a copula based joint framework that ties the collision type (represented as a multinomial logit model) and injury severity (represented as an ordered logit model) through a closed form flexible dependency structure to study the injury severity process. The proposed approach also accommodates the potential heterogeneity (across drivers) in the dependency structure. Further, the study incorporates collision type as a vehicle-level, as opposed to a crash-level variable as hitherto assumed in earlier research, while also examining the impact of a comprehensive set of exogenous factors on driver injury severity. The proposed modeling system is estimated using collision data from the province of Victoria, Australia for the years 2006 through 2010.

1. Background

According to the World Health Organization (WHO), road traffic crashes are one of the major causes of death in the world (WHO, 2013). The economic and societal cost, of road traffic crashes, accrues to billions of dollars (WHO, 2013). For example, in Australia, the total cost of motor vehicle crashes is estimated at approximately $18 billion per annum (Risbey et al., 2010). While improving road infrastructure design to reduce the occurrence of these crashes is essential, it is also important to provide solutions to reduce the consequences in the unfortunate event of a traffic crash. A critical component of identifying and gaining a comprehensive understanding of the factors that contribute to crash outcomes is the estimation and application of disaggregate level crash severity models.

The commonly available traffic crash databases compile injury severity data as an ordinal discrete variable (for example: no injury, minor injury, severe injury, and fatal injury). Naturally, many safety research studies have employed logistic regression[4] approaches (Conroy et al., 2008; Fredette et al., 2008) and ordered discrete outcome models to identify the contributing factors of crash severity (see Savolainen et al., 2011; Yasmin and Eluru, 2013 for a review). Researchers have also employed unordered response models that allow the impact of exogenous variables to vary across injury severity levels. The most prevalent unordered response structure considered is the multinomial logit model (for examples see Schneider et al., 2009; Ulfarsson and Mannering, 2004). More recently, within the ordered response framework, the generalized ordered logit (GOL) (Terza, 1985; Eluru et al., 2008) that enhances the traditional ordered response models has been employed in several safety research efforts (see Yasmin and Eluru, 2013; Eluru, 2013; Mooradian et al., 2013). These research efforts have studied the impact of various exogenous factors that influence injury severity in traffic crashes (see Yasmin and Eluru, 2013 for a detailed review).

* Corresponding author. Tel.: +1 514 398 6823; fax: +1 514 398 7361.
E-mail addresses: shamsunnahar.yasmin@mail.mcgill.ca (S. Yasmin), naveen.eluru@mcgill.ca (N. Eluru), apinjari@eng.usf.edu (A.R. Pinjari), r.tay@latrobe.edu.au (R. Tay).
[1] Tel.: +1 514 398 6823; fax: +1 514 398 7361.
[2] Tel.: +1 813 974 9671; fax: + 813 974 2957.
[3] Tel.: +61 3 9479 1267; fax: +61 3 9479 3283.

[4] To be sure, the logistic regression with two alternatives can be regarded as an ordered logit model with two alternatives.

0001-4575/$ – see front matter © 2014 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.aap.2014.01.018

Most of these studies highlight the collision type variable as one of the most important determinants of vehicle occupant (driver and/or passenger) injury severity. As one would expect, the collision type, whether it is a head-on or a sideswipe, has significant implications for injury severity sustained. For example, the greater dissipation of kinetic energy associated with a head-on collision is likely to result in severe injuries compared to a side-swipe crash. Most of the earlier studies define the collision type as a crash level variable (rear-end, sideswipe, angular, and head-on) – by assigning one collision type for all vehicles involved in the same collision. But, depending on the initial point of impact it is possible that the different vehicles involved in the same crash might have significantly different crash profiles. For example, in a rear-end collision involving two vehicles, one of the vehicle will be rear-ended and the other one will be the rear-ender. The driver of the rear-ended vehicle is likely to be pushed backward into the seat when struck by the rear-ender vehicle leading to a high probability of whiplash or neck injury due to the continuous movement of the neck at a different speed relative to the head and the rest of the body (Khattak, 2001; Chiou et al., 2013; Nordhoff, 2005). Due to the biomechanics of this type of crash, the driver in the rear-ended vehicle is likely to be more seriously injured in a rear-end crash compared to the driver in the rear-ender vehicle. Hence, it is incorrect to assign the same collision type variable to all vehicles involved in the same crash in analyzing vehicle occupant injury severity.[5] The *first contribution* of our research is to address this inconsistency and define a vehicle level collision type variable using a combination of collision type and the initial point of contact.

Most of the earlier studies consider the collision type as an explanatory variable in modeling injury severity (except Ye et al., 2008; Rana et al., 2010). In this approach, the analyst imposes the assumption that the injury severity profile for vehicle occupants in all types of crashes is the same and any potential differences between different collision types can be accurately captured by employing the collision type variable as an explanatory variable. However, it is possible that various collision types might lead to distinct vehicle occupant injury severity profiles i.e., the overall manifestation of injury severity is different by collision type. For example, consider the impact of the gender variable in injury severity models. It is possible that males due to their higher physiological strength are more equipped to resist severe injuries in crashes. However, in a head-on crash due to the greater dissipation of kinetic energy, the physiological advantage might be inadequate. At the same time, the additional strength might be beneficial for male occupants to avoid severe injury in the event of other collision types such as side-swipe. This is an example of how a collision type variable moderates the impact of gender. It is plausible to visualize that collision type variables might similarly affect multiple exogenous variables – indicating that the injury severity profile itself is moderated by the collision type. Thus, estimating a single injury severity model, when such distinct profiles of injury severity exist, will result in incorrect and biased estimates. In fact, several studies have recognized this in safety literature and estimated injury severity focused on a specific type of collision – *Head-on collision*: Gårder, 2006; Conroy et al., 2008; Zuxuan et al., 2006; Zhang and Ivan, 2005; *Rear-end collision*: Khattak, 2001; Yan et al., 2005; Das and Abdel-Aty, 2011; Abdel-Aty and Abdelwahab, 2003; and *Angular collision*: Jin et al., 2010; Chipman, 2004. These studies provide evidence that collision type has a fundamentally distinct effect on injury severity sustained in the crash.

Given the possibility of distinct injury severity profiles – the estimation of separate injury severity models for various collision types seems the appropriate solution. At the same time, it is also important to investigate the factors that result in crashes of a particular collision type. This necessitates a model for collision type; an unordered decision variable that can be studied using a multinomial logit model. Within this system, it is possible that the collision type and resulting injury severity are influenced by the same set of observed and unobserved factors. Accommodating for the impact of observed factors is relatively straightforward within the traditional discrete outcome models by estimating distinct outcome models for collision type (multinomial logit) and injury severity (ordered logit). The process of incorporating the impact of unobserved factors poses methodological challenges. Essentially, accommodating the impact of unobserved factors recognizes that the two dimensions of interest are realizations from the same joint distribution. Traditionally, in econometric literature, such joint processes are examined using simulation based approaches that stitch together the processes through common unobserved error terms (see Eluru and Bhat 2007; Abay et al., 2013 for examples in safety literature). In this direction, Ye et al. (2008) propose a simulation based simultaneous equation framework to study the collision type and injury severity dimensions. The framework employs maximum simulated likelihood approach and requires simulation in the order of the dimension of collision type variables. For instance, in our empirical context, if we have eight vehicle level collision types, it would require us to estimate at least an eight dimensional integral to accommodate for such potential correlations. The process of applying simulation for such joint processes is likely to be error-prone in model estimation as well as inference – particularly the estimation of standard errors (see Bhat, 2011 for a discussion). At the same time, ignoring the presence of such potential jointness may lead to biased and inconsistent parameter estimates in modeling injury severity outcome (Chamberlain, 1980; Eluru and Bhat, 2007; Washington et al., 2003).

More recently, a closed form approach that obviates the need for simulation has been proposed in transportation literature for examining joint decision processes. The approach, referred to as Copula Approach, allows for flexible dependency structures across joint dimensions while retaining the closed form structure (see Bhat and Eluru, 2009). In fact, Rana et al. (2010) employed a copula based approach to consider the crash type and injury severity as a joint process with success. However, both of these studies (Ye et al., 2008; Rana et al., 2010) that jointly model the collision type and injury severity outcome describe the collision type as a crash level variable. But, depending on the position of driver and the initial point of impact, it is possible that the individual vehicle might have different effects in the manner of collision for the same type of collision (see Khattak, 2001 for a discussion in the context of rear-end collision). The *second contribution* of our study is to develop a closed form copula based framework to accommodate the impact of observed and unobserved effects on collision type and injury severity while generating collision type as a vehicle level variable.

The current study enhances the copula based methodology employed by Rana et al. (2010) to study collision type and injury severity. The earlier approach considers the dependency parameter in the copula model to be the same across the entire crash database. However, it is possible that several exogenous factors might actually affect the dependency profile. In other words, the correlation between collision type and injury severity might be stronger or weaker depending on the various attributes of the particular crash. Allowing for such flexibility in the dependency profile allows for more accurate model estimation. The proposed copula dependency parameterization is analogous to the covariance heterogeneity parameterization employed in nested logit models (Bhat, 1997). Ignoring such heterogeneity (when present) will lead to biased

[5] To be sure, Abdel-Aty and Abdelwahab (2003) examine the crash occurrence and Khattak (2001) examine the driver injury outcome by considering the collision type as a vehicle level variable, but only for rear-end collision.

and inconsistent estimates (Chamberlain, 1980; Bhat, 1997). Earlier research efforts have recognized the advantage of such dependency parameterization within the copula framework (see Eluru et al., 2010; Sener et al., 2010). However, these approaches are proposed in the context of joint ordered response structures whereas our study incorporates parameterization of dependency profile in an unordered and ordered joint structure. Our *third contribution* is to formulate the copula model to allow for such potential heterogeneity (across drivers).

The proposed model is estimated using driver injury severity data for two vehicle crashes from the state of Victoria, Australia employing a comprehensive set of exogenous variables – driver characteristics, vehicle characteristics, roadway design attributes, environmental factors and crash characteristics. In summary, the current research effort contributes to safety literature on driver injury severity both methodologically and empirically. *In terms of methodology*, we formulate and estimate a copula-based MNL-OL framework to jointly analyze the collision type and injury severity outcome in a two-vehicle crash. Our study also accommodates the potential heterogeneity (across drivers) in the dependency effect of collision type and injury severity outcome within a closed form copula framework. *In terms of empirical analysis*, our study incorporates collision type as a vehicle level variable and addresses the inconsistency from earlier research while also examining the impact of a comprehensive set of exogenous variables on driver injury severity.

The rest of the paper is organized as follows. Section 2 provides details of the econometric model framework used in the analysis. In Section 3, the data source and sample formation procedures are described. The model results and elasticity effects are presented in Section 4. Section 5 concludes the paper and presents directions for future research.

2. Model framework

The focus of our study is to jointly model the collision type and injury severity outcome of drivers involved in a two vehicle collisions using a copula-based joint multinomial logit-ordered logit modeling framework. The analysis in this paper focuses on driver injury severity in a crash. In this section, econometric formulation for the joint model is presented.

2.1. The collision type outcome model component

Let $q \quad (q = 1, 2, \ldots, Q)$ and $k \quad (k = 1, 2, \ldots, K)$ be the indices to represent driver and collision type, respectively. Let j be the index for the discrete outcome that corresponds to the injury severity level $j \quad (j = 1, 2, \ldots, J)$ of driver q. In the joint framework, the modeling of collision type is undertaken using the multinomial logit structure. Thus, the propensity of a driver q involving in a collision of specific collision type k takes the form of:

$$u^*_{qk} = \beta_k x_{qk} + \xi_{qk} \tag{1}$$

where x_{qk} is a column vector of exogenous variable, β_k is a row vector of unknown parameters specific to collision type k and ξ_{qk} is an idiosyncratic error term (assumed to be standard type-I extreme value distributed) capturing the effects of unobserved factors on the propensity associated with collision type k. A driver q is assumed to be involved in a collision type k if and only if the following condition holds:

$$u^*_{qk} > \max_{l=1,2,\ldots,k,\ \ l \neq k} u^*_{ql} \tag{2}$$

The condition presented in Eq. (2) can be equivalently represented as a series of binary outcome models for each collision type, k (see Lee, 1983). For example, let η_{qk} be a dichotomous variable with $\eta_{qk} = 1$ if a driver q ends up in a collision type k and $\eta_{qk} = 0$ otherwise. Now, let us define v_{qk} as follows[6]:

$$v_{qk} = \xi_{qk} - \left\{ \max_{l=1,2,\ldots,k,\ \ l \neq k} u^*_{ql} \right\} \tag{3}$$

By substituting the right side for u^*_{qk} from Eq. (1) in Eq. (2) we can write:

$$\eta_{qk} = 1 \quad \text{if} \quad \beta'_k x_{qk} + v_{qk} > 0 \tag{4}$$

The system in Eq. (4) represents the multinomial discrete outcome model of collision type as an equivalent series of binary outcome model formulation, one for each collision type k. In Eq. (4), the probability expression of collision type outcome is dependent on the distributional assumption of v_{qk}, which in turn depends on the distributional assumption of ξ_{qk}. Thus an assumption of independent and identical Type 1 Gumbel distribution for ξ_{qk} results in a logistic distributed v_{qk}. Consequently, the probability expression for the corresponding discrete outcome (collision type) model resembles the multinomial logit probability expression as follows:

$$\Lambda_k(\beta_k x_{qk}) = Pr(v_{qk} > -\beta_k x_{qk}) = \frac{\sum_{l \neq k} \exp(\beta_k x_{ql})}{\exp(\beta_k x_{qk}) + \sum_{l \neq k} \exp(\beta_k x_{ql})} \tag{5}$$

2.2. The injury severity outcome model component

In the joint model framework, the modeling of driver injury severity is undertaken using an ordered logit specification. In the ordered response model, the discrete injury severity levels (y_{qk}) are assumed to be associated with an underlying continuous latent variable (y^*_{qk}). This latent variable is typically specified as the following linear function:

$$y^*_{qk} = \alpha_k z_{qk} + \varepsilon_{qk}, \quad y_{qk} = j_k, \quad \text{if} \quad \tau_{k,j-1} < y^*_{qk} < \tau_{k,j} \tag{6}$$

where y^*_{qk} is the latent injury risk propensity for driver q if he/she was involved in a collision type k, z_{qk} is a vector of exogenous variables, α_k is a row vector of unknown parameters and ε_{qk} is a random disturbance term assumed to be standard logistic. $\tau_{k,j}$ ($\tau_{k,0} = -\infty$, $\tau_{k,J} = \infty$) represents the threshold associated with severity level j for collision type k, with the following ordering conditions: $(-\infty < \tau_{k,1} < \tau_{k,2} < \cdots < \tau_{k,J-1} < +\infty)$. Given these relationships across the different parameters, the resulting probability expressions for driver q sustaining an injury severity level j in a collision type k take the following form:

$$Pr(y_{qk} = j_k) = \Lambda_k(\tau_{k,j} - \alpha'_k z_{qk}) - \Lambda_k(\tau_{k,j-1} - \alpha'_k z_{qk}) \tag{7}$$

where $\Lambda_k(\cdot)$ is the standard logistic cumulative distribution function. The probability expression of Eq. (7) represents the independent injury severity model for a collision type k.

[6] The reader would note that the v_{qk} term applied here is different from the Lee's transformation. If one uses a symmetric distribution, that allows both positive and negative dependencies (such as the Gaussian copula proposed by Lee), then Lee's formulation would be adequate. However, when testing various copulas, some of which allow asymmetric and only positive dependencies, it is important to test our version as well as Lee's formulation to ensure we capture the dependencies in asymmetric copulas. We formulate the model in this form because we expect that the dependency for collision type and subsequent injury to be positively correlated (due to unobserved factors, see Portoghese et al., 2011 for a similar formulation in a different context).

2.3. The joint model: a copula-based approach

The collision type and the injury severity component discussed in previous two subsections may be brought together in the following equation system:

$$\eta_{qk} = 1 \quad \text{if} \quad \beta_k x_{qk} > \upsilon_{qk}$$
$$y^*_{qk} = \alpha_k z_{qk} + \varepsilon_{qk}, \quad y_{qk} = 1[\eta_{qk} = 1]y^*_{qk} \tag{8}$$

However, the level of dependency between the underlying collision type outcome and the injury severity level of driver depends on the type and extent of dependency between the stochastic terms υ_{qk} and ε_{qk}. These dependencies (or correlations) are explored in the current study by using a copula-based approach. A copula is a mathematical device that identifies dependency among random variables with pre-specified marginal distribution (Bhat and Eluru (2009) and Trivedi and Zimmer (2007) provide a detailed description of the copula approach). In constructing the copula dependency, the random variables (υ_{qk} and ε_{qk}) are transformed into uniform distributions by using their inverse cumulative distribution functions, which are then coupled or linked as a multivariate joint distribution function by applying the copula structure. Let us assume that $\Lambda_{\upsilon k}(\cdot)$ and $\Lambda_{\varepsilon k}(\cdot)$ are the marginal distribution of υ_{qk} and ε_{qk}, respectively and $\Lambda_{\upsilon k, sk}(\cdot,\cdot)$ is the joint distribution of υ_{qk} and η_{qk}. Subsequently, a bivariate distribution $\Lambda_{\upsilon k,\varepsilon k}(\upsilon, \varepsilon)$ can be generated as a joint cumulative probability distribution of uniform [0, 1] marginal variables U_1 and U_2 as below:

$$\Lambda_{\upsilon k,\varepsilon k}(\upsilon, \varepsilon) = Pr(\upsilon_{qk} < \upsilon, , \varepsilon_{qk} < \varepsilon) = [\Lambda^{-1}_{\upsilon k}(U_1) < \upsilon, \Lambda^{-1}_{\varepsilon k}(U_2) < \varepsilon]$$
$$= [U_1 < \Lambda_{\upsilon k}(\upsilon), U_2 < \Lambda_{\varepsilon k}(\varepsilon)] \tag{9}$$

The joint distribution (of uniform marginal variable) in Eq. (9) can be generated by a function $C_{\theta q}(\cdot, \cdot)$ (Sklar, 1973), such that:

$$\Lambda_{\upsilon k,\varepsilon k}(\upsilon, \delta_2) = C_{\theta q}(U_1 = \Lambda_{\upsilon k}(\upsilon), U_2 = \Lambda_{\varepsilon k}(\varepsilon)) \tag{10}$$

where $C_{\theta q}(\cdot, \cdot)$ is a copula function and θ_q the dependence parameter defining the link between υ_{qk} and ε_{qk}. It is important to note here that, the level of dependence between collision type and injury severity level can vary across drivers. Therefore, in the current study, the dependence parameter θ_q is parameterized as a function of observed crash attributes as follows:

$$\theta_q = fn(\gamma_k s_{qk}) \tag{11}$$

where, s_{qk} is a column vector of exogenous variable, γ_k is a row vector of unknown parameters (including a constant) specific to collision type k and fn represents the functional form of parameterization. Based on the dependency parameter permissible ranges, alternate parameterization forms for the six copulas are considered in our analysis. For Normal, Farlie–Gumbel–Morgenstern (FGM) and Frank Copulas we use $\theta_q = \gamma_k s_{qk}$, for the Clayton copula we employ $\theta_q = \exp(\gamma_k s_{qk})$, and for Joe and Gumbel copulas we employ $\theta_q = 1 + \exp(\gamma_k s_{qk})$.

2.4. Estimation procedure

The joint probability that the driver q gets involved in a collision type k and sustaining injury severity level j, from Eqs. (5) and (7), can be written as:

$$Pr(\eta_{qk} = 1, y_{qk} = j_k) = Pr\{(\beta_k x_{qk} > \upsilon_{qk}), ((\tau_{k,j-1} - \alpha_k z_{qk}) < \varepsilon_{qk}$$
$$< (\tau_{k,j} - \alpha_k z_{qk}))\} = Pr((\beta_k x_{qk} > \upsilon_{qk}), (\varepsilon_{qk} < \tau_{k,j} - \alpha_k z_{qk}))$$
$$- Pr((\beta_k x_{qk} > \upsilon_{qk}), (\varepsilon_{qk} < \tau_{k,j-1} - \alpha_k z_{qk}))\Lambda_{\varepsilon k}(\tau_{k,j} - \alpha_k z_{qk})$$
$$- \Lambda_{\varepsilon k}(\tau_{k,j-1} - \alpha_k z_{qk}) - (Pr[\upsilon_{qk} < -\beta_k x_{qk}, \varepsilon_{qk} < (\tau_{k,j} - \alpha_k z_{qk})$$
$$- Pr[\upsilon_{qk} < -\beta_k x_{qk}, \varepsilon_{qk} < (\tau_{k,j-1} - \alpha_k z_{qk})]) \tag{12}$$

The joint probability of Eq. (12) can be expressed by using the copula function in Eq. (10) as:

$$Pr(\eta_{qk} = 1, y_{qk} = j_k) = \Lambda_{\varepsilon k}(\tau_{k,j} - \alpha_k z_{qk}) - \Lambda_{\varepsilon k}(\tau_{k,j-1} - \alpha_k z_{qk})$$
$$- [C_{\theta q}(U^k_{q,j}, U^k_q) - C_{\theta q}(U^k_{q,j-1}, U^k_q)] \tag{13}$$

where

$$U^k_{q,j} = \Lambda_{\varepsilon k}(\tau_{k,j} - \alpha_k z_{qk}), \quad U^k_q = \Lambda_{\upsilon k}(-\beta_k x_{qk}) \tag{14}$$

Thus the likelihood function with the joint probability expression in Eq. (13) for collision type and driver injury severity outcomes can be expressed as:

$$L = \prod_{q=1}^{Q} \left[\prod_{k=1}^{K} \prod_{j=1}^{J} \{Pr(\eta_{qk} = 1, y_{qk} = j_k)\}^{\omega_{qkj}} \right] \tag{15}$$

where ω_{qkj} is dummy with $\omega_{qkj} = 1$ if the driver q sustains collision type k and an injury severity level of j and 0 otherwise. All the parameters in the model are then consistently estimated by maximizing the logarithmic function of L. The parameters to be estimated in the model are: β_k in the MNL component, α_k and $\tau_{k,j}$ in OL component, and finally γ_k in the dependency component. In our analysis we employ six different copulas structure – the Gaussian copula, the Farlie–Gumbel–Morgenstern (FGM) copula, and set of Archimedean copulas including Frank, Clayton, Joe and Gumbel copulas (a detailed discussion of these copulas is available in Bhat and Eluru, 2009).

3. Data

3.1. Data source

Data for our empirical analysis are sourced from the Victoria crash database of Australia for the years 2006 through 2010. For the five years, the crash database has a record of 67,809 crashes involving 118,842 motor vehicles and 166,040 individuals resulting in 1550 fatalities and 87,855 injuries to the crash victims. A four point ordinal scale is used in the database to represent the injury severity of individuals involved in these crashes: (1) no injury; (2) minor injury; (3) serious injury and (4) fatal injury.

3.2. Sample formation and the dependent variables

This study is confined to the injury severity outcome of drivers, who are involved in a two passenger vehicle collisions. Crashes involving only one vehicle or more than two vehicles are not included in the analysis. The crashes that involve commercial vehicles are also excluded to avoid the potential systematic differences between the crashes involving commercial and non-commercial driver groups.

In our analysis, the crash outcome is defined as the injury severity level sustained by the driver in each vehicle of the two vehicle collisions. The final dataset, after removing records with missing information for essential attributes consisted of about 34,278 driver records. In this final sample of drivers, the percentage of fatal crashes sustained by drivers is extremely small (0.40%). Therefore, both the fatal and serious injury categories are merged together. From this dataset, a sample of 8509 driver records is randomly selected for the purpose of estimating models. In the final estimation sample, the distributions of the *three* driver injury severity levels are as follows: no injury 49.50%, minor injury 34.50% and serious/fatal injury 16.00%.

As discussed earlier, the database compiles the types of collision at a high level of disaggregation, and as a combination of collision type (rear-end, sideswipe, angular, and head-on) and the initial point of contact.[7]

A schematic diagram of the initial point of impact relative to the driver's seat position is shown in Fig. 1. Based on the collision type and the point of impact, we identified *eight categories* for the "collision type": rear-ender (the rear vehicle that is involved in rear-end collision), rear-ended (the front vehicle that is involved in the rear-end collision), near-sideswipe (sideswipe/near-side), far-sideswipe (sideswipe/far-side), near-angular (angular/near-side), far-angular (angular/far-side), short-side angular (angular/front and rear side) and head-on (head-on/front side). In the final estimation sample, the distribution of collision type variable is as follows: rear-ender 11.91% rear-ended 14.29%, near-sideswipe 2.49%, far-sideswipe 3.04%, near-angular 17.95%, far-angular 16.61%, short-side angular 26.80% and head-on 6.92%.

Table 1 offers a summary of the sample characteristics of collision type and injury severity level sustained by drivers. From the descriptive analysis, it is evident that the injury severity distributions vary substantially by collision type. More interestingly, we observe that for collision types within the same accident, rear-ender vs. rear-ended, near-sideswipe vs. far-sideswipe exhibit huge differences in the injury severity distribution. These observations highlight the need to define the collision type variable at a vehicle level rather than at the crash level. The descriptive analysis identifies head-on as the most serious collision type in terms of severe injuries while far-sideswipe crashes result in the least severe injuries. Further, Table 2 offers a summary of the sample characteristics of explanatory variables across different collision types. It can be observed from Table 2 that the proportions of different variables vary substantially across different collision types.

4. Empirical analysis

4.1. Variables considered

The collision attributes considered in the empirical study can be grouped into the following five broad categories:

- Driver characteristics including driver age, gender, seat belt use and local driver information;
- Vehicle characteristics including vehicle type (characterized as sedan, station wagon, utility and panel van) and vehicle age;
- Roadway design attributes including type of road surface, presence of traffic control device, speed zones and type of intersection;
- Environmental factors including time of day, day of week, weather condition, surface condition and lighting condition; and
- Crash characteristics including presence of passenger and trajectory of vehicle's motion.

The final specification of the model development was based on combining the variables when their effects were not statistically different and by removing the statistically insignificant variables in a systematic process based on statistical significance (90% confidence level). The coefficient estimates across different collision types were also restricted to be same when the effects were not significantly different.

4.2. Model specification and overall measures of fit

The empirical analysis involves estimation of models by using six different copula structures: (1) Gaussian, (2) FGM, 3) Clayton, (4) Gumbel, (5) Frank and (6) Joe (a detailed discussion of these copulas is available in Bhat and Eluru, 2009). The empirical analysis involved a series of model estimations. *First*, an independent copula model (separate MNL and OL models) were estimated to establish a benchmark for comparison. *Second*, 6 different models that restricted the dependency parameters across the eight collision types and injury severity models to be the same were estimated. *Third*, based on the copula parameter significance for each collision type, copula models that allow for different dependency structures for different collision type and injury severity combinations were estimated (for example Frank copula for the first three collision types Clayton copula for other collision types). Finally, to determine the most suitable copula model (including the independent copula model), a comparison exercise was undertaken. The alternative copula models estimated are non-nested and hence, cannot be tested using traditional log-likelihood ratio test. We employ the Bayesian Information Criterion (BIC) to determine the best model among all copula models (see Trivedi and Zimmer, 2007; Quinn, 2007; Eluru et al., 2010). The BIC for a given empirical model is equal to:

$$BIC = -2LL + K\ \ln(Q) \tag{16}$$

where LL is the log-likelihood value at convergence, K is the number of parameters, and Q is the number of observations. The model with the *lower* BIC is the preferred copula model. With exclusively a single copula dependency structure, the best model fit is obtained with Clayton. However, the lowest BIC value was obtained for a combination model of Frank–Clayton copulas (Frank copula structure for rear-ender and head-on collision and Clayton dependency structure with the remaining collision type). The copula model BIC comparisons confirm the importance of accommodating dependence between collision type and injury severity outcome in the analysis of driver injury severity.

4.3. Estimation results

In presenting the effects of exogenous variables in the joint model specification, we will restrict ourselves to the discussion of the Frank–Clayton specification. For the ease of presentation, the collision type component (Table 3) and injury severity component (Table 4) are presented and discussed separately. The copula parameters are presented in the last row panel of Table 3.

4.3.1. Collision type component

The coefficients in Table 3 represent the effect of exogenous variables on each collision type category relative to the base category. In the following sections, the estimation results are discussed by variable groups.

Driver characteristics: The impact of driver age on collision type indicates that young drivers are more likely to be the rear-ender and are less likely to be rear-ended in crashes relative to the adult drivers, perhaps reflecting a lack of driving experience and/or poor judgment and/or a greater risk-taking/aggressive driving propensity. The likelihood of being rear-ended or being involved in a far-sideswipe collision is lower for the older drivers. However, the older drivers are also more likely to be involved in angular collision (far- or near-angular) compared to the adult drivers, which might be a manifestation of longer time requirements for older drivers in complete turning movements (Alexander et al., 2002). Female

[7] It is worthwhile to mention here that several previous studies (Tsui et al., 2009; Schiff and Cummings, 2004; Loo and Tsui, 2007) have examined the reliability of crash related factors documented in police-reported crash databases. The unreliability in reporting is mostly observed for casualty of crash, occupant position in the vehicle, demographics and seat-belt. Compiling crash details based on collision type and initial point of impact are less likely to be error prone. More importantly, the incompleteness of these variables in the Victorian crash database is approximately zero (zero for collision type and 0.3% for initial point of impact).

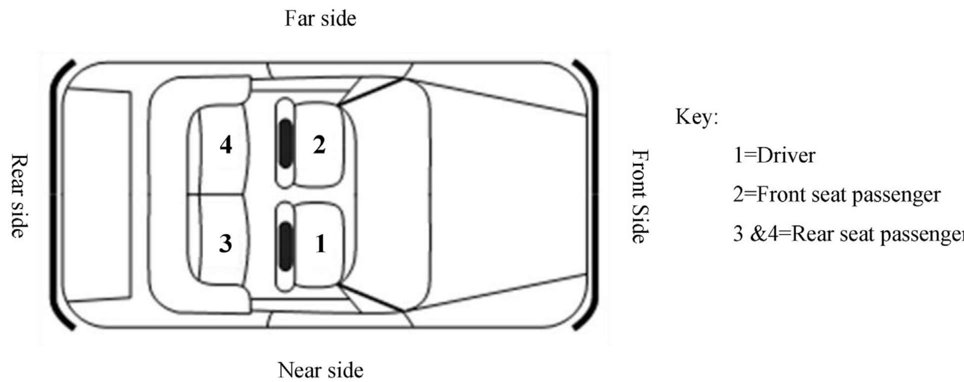

Fig. 1. Schematic diagram of initial point of impact relative to the drivers' seat position.

drivers are more likely to be rear-ended or involved in a near-angular collision, while the odds of involving in a head-on collision is lower for female drivers compared to their male counterparts. The results also highlight that drivers who do not wear seat-belts are more likely to hit another vehicle from behind, a possible reflection of inherent aggressive personality of these drivers.

Vehicle characteristics: The effects of the vehicle characteristics indicate that the drivers of utility and panel van are more likely to be the rear-ender, while the likelihood of being involved in head-on collisions are also higher for the driver of utility vehicle compared to other drivers. These results point toward aggressive attitude in driving and a false sense of security among large vehicle owners. The vehicle age variables suggest that compared to the drivers of newer vehicles (vehicle age less than 6), the drivers of older vehicles (vehicle age 6–10 or vehicle age 11 and above) are less likely to be rear-ended or involved in any form of sideswipe or angular collision (the effect of vehicle age 6–10 is insignificant for far-sideswipe collision).

Roadway design attributes: Among the roadway design attributes, the effect of roadway surface type is significant only for the head-on collision with positive coefficient for the gravel road surface compared to the paved and unpaved roads. Usually, gravel roads are associated with fewer lanes increasing the odds of head-on collisions as the lanes are unlikely to be median separated. The estimation results corresponding to the presence of traffic control device highlight that the presence of traffic signal is associated with less sideswipe and head-on collision. Drivers are more likely to be rear-ended or involved in near-angular collision in the presence of roundabout. In the presence of a stop sign, the likelihood of rear-end, far-sideswipe and head-on collisions are lower, whereas the likelihood of near-angular collision is higher. The presence of yield sign has positive association with rear-ended and near-angular collision and negative association with sideswipe and head-on collision.

With respect to the speed zone, the medium speed limit zone indicator increases the likelihood of rear-end collision; while the high speed limit zone indicators reveal increased likelihood of rear-ended, side-swipe and head-on collisions. The presence of T-intersection increases the odds of all collision types (except short-side angular). Five or more legged intersection is positively correlated with the occurrence of rear-end and far-sideswipe collision. The variable representing the location as a non-intersection is associated with higher crash propensity for all collision types except for far- and short side-angular collision.

Environmental factors: The effects of environmental factors indicate that the occurrence of far-angular collision is less at late night compared to the other times of day. Crashes occurring on wet surface are more likely to be head-on collision than those occurring on the dry surface condition. Far-sideswipe collision is less likely to occur at dawn/dusk period relative to the daylight period. Dark-lighted condition results in reduced likelihood of rear-end collision. However, dark-unlighted condition is associated with high risk of head-on collision. During weekend, drivers are less likely to be involved in rear-ended situations, but are more likely to be involved in far-sideswipe and head-on collisions.

Crash characteristics: Among the crash characteristic variables considered, none of the variables show significant impact on collision type occurrence.

4.3.2. Dependence effects

As indicated earlier, the estimated Frank–Clayton copula based MNL-OL model provides the best fit in incorporating the correlation between the collision type and injury severity outcome. An examination of the copula parameters presented in the last row panel of Table 3 highlights the presence of common unobserved factors affecting collision type and injury severity. The Frank copula dependency structure is associated with the rear-ender and head-on collision types, while the Clayton dependency structure is associated with the rest of the six collision types. Further, except for far-angular collision type, all other copula dependencies are characterized by at least one additional exogenous variable. This provides support to our hypothesis that the dependency structures

Table 1
Sample characteristics of collision type and injury severity level sustained by drivers.

Injury severity	Collision type							
	Rear-ender	Rear-ended	Near-sideswipe	Far-sideswipe	Near-angular	Far-angular	Short-side angular	Head-on
No injury	612 (60.41%)[a]	422 (34.70%)	101 (47.64%)	183 (70.66%)	701 (45.91%)	803 (56.83%)	1211 (53.11%)	175 (29.71%)
Minor injury	261 (25.77%)	659 (54.19%)	72 (33.96%)	59 (22.78%)	526 (34.45%)	432 (30.57%)	718 (31.49%)	210 (35.65%)
Serious/fatal injury	140 (13.82%)	135 (11.10%)	39 (18.40%)	17 (6.56%)	300 (19.65%)	178 (12.60%)	351 (15.39%)	204 (34.63%)
Total	1013	1216	212	259	1527	1413	2280	589

[a] The numbers in parenthesis correspond to column percentages.

Table 2
Sample characteristics of explanatory variables across different collision types.

Variables	Rear-ender	Rear-ended	Near-sideswipe	Far-sideswipe	Near-angular	Far-angular	Short-side angular	Head-on
Driver characteristics								
Driver age								
Age less than 25	306(30.21)[a]	212(17.43)	57(26.89)	55(21.24)	343(22.46)	332(23.50)	556(24.39)	131(22.24)
Age 25–64	612(60.81)	913(75.09)	133(62.73)	190(26.65)	952(62.35)	885(62.63)	1473(64.6)	401(68.08)
Age above 65+	91(8.98)	91(7.48)	22(10.38)	14(5.41)	232(15.19)	196(13.87)	251(11.01)	57(9.68)
Driver gender								
Female	439(43.34)	693(56.99)	92(43.40)	113(43.63)	790(51.74)	648(45.86)	1049(46.01)	205(34.80)
Male	574(56.66)	523(43.01)	120(56.60)	146(56.37)	737(48.26)	765(54.14)	1231(53.99)	384(65.20)
Restraint system use								
Seat belt not used	38(3.75)	32(2.63)	9(4.25)	11(4.25)	35(2.29)	31(2.19)	61(2.68)	22(3.74)
Seat belt used	975(96.25)	1184(97.37)	203(95.75)	248(95.75)	1492(97.71)	1382(97.81)	2219(97.32)	567(96.26)
Locality of driver								
Non-local driver	119(11.75)	147(12.09)	33(15.57)	36(13.90)	145(9.50)	121(8.56)	199(8.73)	109(18.51)
Local driver	894(88.25)	1069(87.91)	179(84.43)	223(86.10)	1382(90.50)	1292(91.44)	2081(91.27)	480(81.49)
Vehicle characteristics								
Vehicle type								
Car	688(67.92)	887(72.94)	154(72.64)	182(70.27)	1099(71.97)	1013(71.69)	1684(73.86)	378(64.18)
Station wagon	177(17.47)	219(18.01)	34(16.04)	50(19.31)	285(18.66)	248(17.55)	395(17.32)	118(20.03)
Utility	108(10.66)	85(6.99)	17(8.02)	21(8.11)	108(7.07)	118(8.35)	159(6.97)	80(13.58)
Panel van	40(3.95)	25(2.06)	7(3.30)	6(2.32)	35(2.29)	34(2.41)	42(1.84)	13(2.21)
Vehicle age								
Vehicle age less than 6	282(27.84)	404(33.22)	75(35.38)	88(33.98)	496(32.48)	439(31.07)	636(27.89)	172(29.20)
Vehicle age 6–10	297(29.32)	333(27.38)	57(26.89)	87(33.59)	373(24.43)	383(27.11)	651(28.55)	158(26.83)
Vehicle age 11 and above	434(42.84)	479(39.39)	80(37.74)	84(32.43)	658(43.09)	591(41.83)	993(43.55)	259(43.97)
Roadway design attributes								
Type of road surface (base: paved)								
Paved	990(97.73)	1189(97.78)	206(97.17)	254(98.07)	1483(97.12)	1385(98.02)	2231(97.85)	531(90.15)
Unpaved	1(0.10)	1(0.08)	0(0.00)	1(0.39)	1(0.07)	0(0.00)	2(0.09)	6(1.02)
Gravel	22(2.17)	26(2.14)	6(2.83)	4(1.54)	43(2.82)	28(1.98)	47(2.06)	52(8.83)
Traffic control device								
No control	622(61.40)	730(60.03)	177(83.49)	211(81.47)	587(38.44)	638(45.15)	957(41.97)	566(96.10)
Signal	237(23.40)	304(25.00)	23(10.85)	34(13.13)	319(20.89)	437(30.93)	764(33.51)	4(0.68)
Other traffic control	13(1.28)	29(2.38)	0(0.00)	4(1.54)	26(1.70)	24(1.70)	40(1.75)	8(1.36)
Pedestrian control	11(1.09)	12(0.99)	1(0.47)	0(0.00)	4(0.26)	2(0.14)	7(0.31)	0(0.00)
Roundabout	30(2.96)	44(3.62)	8(3.77)	6(2.32)	70(4.58)	65(4.60)	86(3.77)	4(0.68)
Stop sign	14(1.38)	13(1.07)	0(0.00)	1(0.39)	157(10.28)	60(4.25)	125(5.48)	1(0.17)
Yield sign	86(8.49)	84(6.91)	3(1.42)	3(1.16)	364(23.84)	187(13.23)	301(13.20)	6(1.02)
Speed zone								
Low speed (≤50 km/h)	118(11.65)	141(11.60)	30(14.15)	42(16.22)	326(21.35)	298(21.09)	461(20.22)	74(12.56)
Medium speed (60–90 km/h)	783(77.30)	952(78.29)	141(66.51)	164(63.32)	1057(69.22)	1016(71.90)	1664(72.98)	314(53.31)
High speed (≥100 km/h)	112(11.06)	123(10.12)	41(19.34)	53(20.46)	144(9.43)	99(7.01)	155(6.80)	201(34.13)
Type of intersection								
Cross intersection	282(27.84)	293(24.10)	34(16.04)	43(16.60)	654(42.83)	701(49.61)	1108(48.60)	16(2.72)
T intersection	263(25.96)	361(29.69)	52(24.53)	66(25.48)	569(37.26)	471(33.33)	832(36.49)	69(11.71)
Y intersection	5(0.49)	4(0.33)	0(0.00)	0(0.00)	6(0.39)	6(0.42)	10(0.44)	2(0.34)
Five and more legged intersection	28(2.76)	43(3.54)	2(0.94)	9(3.47)	39(2.55)	43(3.04)	78(3.42)	1(0.17)
Non-intersection	435(42.94)	515(42.35)	124(58.49)	141(54.44)	259(16.96)	192(13.59)	251(11.01)	501(85.06)

Environmental factors								
Time of day								
Morning peak	138 (13.62)	168 (13.82)	36 (16.98)	44 (16.99)	229 (15.00)	194 (13.73)	311 (13.64)	82 (13.92)
Off peak	358 (35.34)	440 (36.18)	69 (32.55)	97 (37.45)	514 (33.66)	489 (34.61)	774 (33.95)	183 (31.07)
Evening peak	291 (28.73)	352 (28.95)	54 (25.47)	52 (20.08)	419 (27.44)	368 (26.04)	588 (25.79)	148 (25.13)
Late evening	197 (19.45)	233 (19.16)	49 (23.11)	54 (20.85)	324 (21.22)	329 (23.28)	532 (23.33)	144 (24.45)
Late night	29 (2.86)	23 (1.89)	4 (1.89)	12 (4.63)	41 (2.69)	33 (2.34)	75 (3.29)	32 (5.43)
Weather condition								
Clear	863 (85.19)	1058 (87.01)	183 (86.32)	228 (88.03)	1343 (87.95)	1270 (89.88)	1981 (86.89)	446 (75.72)
Rainy/snowy/foggy	131 (12.93)	138 (11.35)	24 (11.32)	26 (10.04)	168 (11.00)	128 (9.06)	271 (11.89)	127 (21.56)
High wind	19 (1.88)	20 (1.64)	5 (2.36)	5 (1.93)	16 (1.05)	15 (1.06)	28 (1.23)	16 (2.72)
Surface condition								
Dry	837 (82.63)	1025 (84.29)	178 (83.96)	226 (87.26)	1297 (84.94)	1230 (87.05)	1906 (83.60)	403 (68.42)
Wet	170 (16.78)	187 (15.38)	34 (16.04)	33 (12.74)	226 (14.80)	182 (12.88)	371 (16.27)	175 (29.71)
Muddy	3 (0.30)	1 (0.08)	0 (0.00)	0 (0.00)	0 (0.00)	1 (0.07)	1 (0.04)	6 (1.02)
Snowy	3 (0.30)	3 (0.25)	0 (0.00)	0 (0.00)	4 (0.26)	0 (0.00)	2 (0.09)	5 (0.85)
Lighting condition								
Day	760 (75.02)	933 (76.73)	153 (72.17)	197 (76.06)	1139 (74.59)	1033 (73.11)	1617 (70.92)	412 (69.95)
Dusk/dawn	79 (7.80)	78 (6.41)	18 (8.49)	10 (3.86)	92 (6.02)	88 (6.23)	166 (7.28)	29 (4.92)
Dark-lighted	137 (13.52)	173 (14.23)	38 (17.92)	43 (16.60)	267 (17.49)	269 (19.04)	446 (19.56)	74 (12.56)
Dark-unlighted	30 (2.96)	27 (2.22)	3 (1.42)	8 (3.09)	27 (1.77)	16 (1.13)	43 (1.89)	72 (12.22)
Other lighting condition	7 (0.69)	5 (0.41)	0 (0.00)	1 (0.39)	2 (0.13)	7 (0.50)	8 (0.35)	2 (0.34)
Days of week								
Weekend	221 (21.82)	222 (18.26)	52 (24.53)	80 (30.89)	351 (22.99)	359 (25.41)	548 (24.04)	198 (33.62)
Weekday	792 (78.18)	994 (81.74)	160 (75.47)	179 (69.11)	1176 (77.01)	1054 (74.59)	1732 (75.96)	391 (66.38)
Crash characteristics								
Trajectory of vehicle's motions								
Going straight	742 (73.25)	311 (25.58)	75 (35.38)	110 (42.47)	687 (44.99)	702 (49.68)	1655 (72.59)	418 (70.97)
Other movement	271 (26.75)	905 (74.42)	137 (64.62)	149 (57.53)	840 (55.01)	711 (50.32)	625 (27.41)	171 (29.03)
Presence of passenger								
No passenger	2 (0.20)	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)	1 (0.17)
One passenger	559 (55.18)	689 (56.66)	121 (57.08)	139 (53.67)	797 (52.19)	676 (47.84)	1081 (47.41)	258 (43.80)
Two passenger	256 (25.27)	305 (25.08)	53 (25.00)	64 (24.71)	363 (23.77)	393 (27.81)	655 (28.73)	157 (26.66)
More than two passengers	196 (19.35)	222 (18.26)	38 (17.92)	56 (21.62)	367 (24.03)	344 (24.35)	544 (23.86)	173 (29.37)

[a] The numbers in parenthesis correspond to column percentages within each category.

Table 3
MNL (collision type) model estimates and copula parameters.

Variables	Rear-ender	Rear-ended	Near-sideswipe	Far-sideswipe	Near-angular	Far-angular	Short-side angular	Head-on
Constant	–	−1.957 (0.090)[a]	−1.617 (0.106)	−2.750 (0.153)	−2.705 (0.132)	−0.846 (0.072)	−0.352 (0.056)	−3.846 (0.276)
Driver characteristics								
Driver age (base: age 25–64)								
Age less than 25	0.361 (0.077)	−0.424 (0.085)	–	–	–	–	–	–
Age above 65+	–	−0.416 (0.125)	–	−0.711 (0.286)	0.311 (0.076)	0.311 (0.076)	–	–
Driver gender (base: male)								
Female	–	0.463 (0.068)	–	–	0.239 (0.060)	–	–	−0.264 (0.106)
Restraint system use (base: seat belt used)								
Seat belt not used	0.378 (0.180)	–	–	–	–	–	–	–
Vehicle characteristics								
Vehicle type (base: sedan)								
Utility	0.310 (0.100)	–	–	–	–	–	–	0.310 (0.100)
Panel van	0.609 (0.184)	–	–	–	–	–	–	–
Vehicle age (base: vehicle age less than 6)								
Vehicle age 6–10	–	−0.181 (0.064)	−0.341 (0.074)	–	−0.341 (0.074)	−0.181 (0.064)	–	–
Vehicle age 11 and above	–	−0.169 (0.076)	−0.391 (0.151)	−0.474 (0.142)	−0.140 (0.058)	−0.140 (0.058)	–	–
Roadway design attributes								
Type of road surface (base: paved)								
Gravel	–	–	–	–	–	–	–	1.361 (0.250)
Traffic control device (base: no traffic control and other control device)								
Signal	–	–	−0.7095 (0.163)	−0.709 (0.163)	–	–	–	−2.8209 (0.547)
Pedestrian control	1.381 (0.356)	1.381 (0.356)	–	–	–	–	–	–
Roundabout	–	0.309 (0.182)	–	–	0.494 (0.148)	–	–	–
Stop sign	−0.894 (0.282)	−1.261 (0.301)	–	−2.406 (1.017)	1.070 (0.118)	–	–	−2.657 (1.162)
Yield sign	–	0.314 (0.126)	−2.094 (0.426)	−2.094 (0.426)	0.937 (0.80)	–	–	−1.692 (0.442)
Speed zone (base: low speed zone ≤50 km/h)								
Medium speed (60–90 km/h)	0.616 (0.073)	0.616 (0.073)	–	–	–	–	–	–
High speed (≥100 km/h)	–	0.575 (0.116)	0.575 (0.116)	0.671 (0.186)	–	–	–	0.934 (0.133)
Type of intersection (base: cross intersection)								
T intersection	0.217 (0.092)	0.438 (0.082)	0.4378 (0.082)	0.634 (0.182)	0.143 (0.075)	−0.136 (0.071)	–	1.384 (0.296)
Five and more legged intersection	0.449 (0.215)	0.708 (0.181)	–	1.184 (0.375)	–	–	–	–
Non-intersection	1.807 (0.077)	1.807 (0.077)	2.141 (0.137)	2.141 (0.137)	0.768 (0.093)	–	–	3.864 (0.277)
Environmental factors								
Time of day (base: morning peak, off peak and late evening)								
Late night	–	–	–	–	–	−0.3816 (0.196)	–	–
Surface condition (base: dry)								
Wet	–	–	–	–	–	–	–	0.933 (0.113)
Lighting condition (base: daylight)								
Dusk/dawn	–	–	–	−0.557 (0.339)	–	–	–	–
Dark-lighted	−0.407 (0.101)	−0.243 (0.091)	–	–	–	–	–	–
Dark-unlighted	–	–	–	–	–	–	–	0.969 (0.176)
Days of week								
Weekend	–	−0.278 (0.083)	–	0.406 (0.145)	–	–	–	0.504 (0.111)

	Copula parameters							
	Frank	Clayton	Clayton	Clayton	Clayton	Clayton	Clayton	Frank
Constant	3.047 (1.667)	1.423 (0.383)	0.495 (0.625)	0.636 (0.602)	0.772 (0.511)	2.661 (0.582)	1.473 (0.414)	1.783 (1.046)
Female driver	0.971 (0.540)	–	–	–	–	–	–	–
Medium speed limit	–	–	–	–	1.482 (0.228)	–	–	0.943 (0.527)
Yield sign	–	1.651 (0.285)	–	–	–	–	–	–
Utility vehicle	–	–	3.978 (0.737)	–	–	–	–	–
Late night	–	–	–	5.079 (0.904)	–	–	–	–
High wind	–	–	–	–	2.783 (0.599)	–	2.318 (0.497)	–

[a] Standard errors are presented in parenthesis.

are not constant across the entire database. The various exogenous variables that contribute to the dependency include Female (rear-ender), medium speed limit (near-angular and head-on), yield sign (rear-ended), utility vehicle (near-sideswipe), late night (far-sideswipe) and high wind (near-angular and short-side angular). The Frank copula offers a symmetric dependency structure i.e. a positive coefficient represents a positive dependency while negative coefficient represents negative dependency. The exact nature of the dependency for the Frank copula is based on the realized coefficient for rear-ender and head-on crash types considering all significant variables. For the Clayton copula, the dependency is entirely positive and the coefficient sign and magnitude reflects whether a variable increases or reduces the dependency and by how much. The proposed framework by allowing for such parameterizations allows us to improve the model estimation results.

4.3.3. *Injury severity component*

The coefficients in Table 4 represent the effect of exogenous variables on injury severity outcome of drivers for each collision type category. The results suggest that the impact of exogenous variables vary (for some variables) in magnitude as well as in sign across collision types. The impacts of these variables are also substantially different from the estimates of independent MNL-OL model (the results are not presented here to conserve on space). For instance, the differences in variable estimates (independent MNL-OL model and copula based MNL-OL model) are more than 20% in *rear-ender* for high wind and T-intersection; in *rear-ended* for high wind, presence of one passenger and two passenger; in far-angular for medium speed limit and high speed limit; in *short-side angular* for medium speed limit and high speed limit; and in *head-on* collision for weekend and morning peak-period.

In the following sections, the estimation results for injury severity component of the joint model are discussed by variable groups.

Driver characteristics: The impacts of driver characteristics reveal significant variations based on driver age, gender, seat-belt use and driver knowledge of local conditions. The results indicate that the likelihood of being severely injured is lower for the young drivers compared to the adult drivers, particularly for rear-ended and short side-angular collisions, perhaps indicating the higher physiological strength of young drivers. Compared to the adult drivers, older drivers are more likely to sustain serious injury across a range of collision type, a result also observed in several previous studies (Bédard et al., 2002; Kim et al., 2013; Williams and Shabanova, 2003). Female drivers are consistently associated with higher injury risk propensity across all collision type presumably because of their lower physiological strength compared to their male counterparts. The negative impact of not using seat-belt is found significant only for near-angular collision type. The driver knowledge of local conditions characterized as local versus non-local drivers reveals that non-local drivers are likely to sustain serious injury for rear-ended, far-sideswipe and near-angular collisions. Driver unfamiliarity with the driving environment and road rules might contribute to severe driver injuries.

Vehicle characteristics: With respect to driver's vehicle type, the results indicate that drivers in station wagon are less likely to be severely injured compared to other drivers for seven of the eight collision types. The finding is consistent with the notion that heavier vehicles provide increased protection to drivers from severe injury. The positive effect of driving larger vehicles is significant in short side-angular and head-on collision for drivers of SUV and panel van. Consistent with several previous studies (Kim et al., 2013; Islam and Mannering, 2006) for most of the collision types, drivers in older vehicles (either vehicle age 6–10 or vehicle age 11 and above) have higher injury risk propensity compared to drivers

Table 4
OL (injury severity) model estimates.

Variables	Rear-ender	Rear-ended	Near-sideswipe	Far-sideswipe	Near-angular	Far-angular	Short-side angular	Head-on
Threshold 1	1.970 (0.406)[a]	1.405 (0.207)	1.405 (0.207)	2.122 (0.192)	1.405 (0.207)	3.010 (0.343)	2.122 (0.192)	0.332 (0.400)
Threshold 2	3.413 (0.347)	4.021 (0.163)	2.947 (0.172)	4.021 (0.163)	2.947 (0.172)	4.593 (0.294)	3.685 (0.164)	1.904 (0.319)
Driver characteristics								
Driver age (base: age 25–64)								
Age less than 25	–	−0.437 (0.131)	–	–	–	–	−0.195 (0.086)	–
Age above 65+	0.454 (0.107)	–	1.182 (0.385)	–	0.569 (0.080)	0.454 (0.107)	0.569 (0.080)	–
Driver gender (base: male)								
Female	0.7714 (0.046)	0.7714 (0.046)	0.7714 (0.046)	0.7714 (0.046)	0.7714 (0.046)	0.7714 (0.046)	0.7714 (0.046)	0.7714 (0.046)
Restraint system use (base: seat belt used)								
Seat belt not used	–	–	–	–	0.616 (0.288)	–	–	–
Locality of driver (base: local driver)								
Non-local driver	–	0.272 (0.103)	–	0.718 (0.384)	0.272 (0.103)	–	–	–
Vehicle characteristics								
Vehicle type (base: sedan)								
Station wagon	−0.4827 (0.071)	−0.237 (0.079)	–	−1.100 (0.508)	−0.483 (0.071)	−0.237 (0.079)	−0.483 (0.071)	−0.237 (0.079)
Utility	–	–	–	–	–	–	−0.690 (0.178)	–
Panel van	–	–	–	–	–	–	–	−0.846 (0.491)
Vehicle age (base: vehicle age less than 6)								
Vehicle age 6–10	0.214 (0.059)	–	0.214 (0.059)	–	–	0.214 (0.059)	0.214 (0.059)	–
Vehicle age 11 and above	0.297 (0.047)	0.297 (0.047)	0.297 (0.047)	0.297 (0.047)	0.297 (0.047)	–	0.297 (0.047)	–
Roadway design attributes								
Type of road surface (base: paved)								
Gravel	–	–	–	–	–	–	–	−0.558 (0.332)
Traffic control device (base: none traffic control and other control device)								
Signal	−0.392 (0.148)	–	–	–	0.228 (0.113)	0.572 (0.101)	–	–
Pedestrian control	–	−0.969 (0.440)	–	–	–	–	−0.969 (0.440)	–
Roundabout	–	–	–	–	−1.227 (0.188)	–	−1.227 (0.188)	–
Stop sign	–	–	–	–	–	–	−0.409 (0.165)	–
Yield sign	−0.942 (0.275)	–	–	–	–	–	−0.317 (0.113)	–
Speed zone (base: low speed zone ≤50 km/h)								
Medium speed (60–90 km/h)	–	–	–	–	–	0.343 (0.117)	0.419 (0.096)	–
High speed (≥100 km/h)	0.844 (0.102)	–	0.844 (0.102)	0.844 (0.102)	0.844 (0.102)	1.187 (0.132)	1.187 (0.132)	0.844 (0.102)
Type of intersection								
T intersection	0.248 (0.137)	–	–	−1.231 (0.423)	–	–	–	–
Five or more legged intersection	−1.007 (0.510)	–	–	–	–	–	–	–
Non-intersection	–	−0.254 (0.079)	–	–	–	–	−0.254 (0.079)	–
Environmental factors								
Time of day (base: morning peak, off peak and late evening)								
Morning peak	–	–	–	–	–	–	–	0.694 (0.211)
Late night	1.202 (0.231)	–	–	–	–	–	0.504 (0.184)	1.202 (0.231)
Weather condition (base: clear)								
Rainy/snowy/foggy	–	–	0.727 (0.139)	–	–	0.727 (0.139)	–	–
High wind	−0.807 (0.339)	−0.807 (0.339)	–	–	–	–	–	–
Lighting condition (base: daylight)								
Dusk/dawn	–	–	–	–	–	–	0.281 (0.141)	0.621 (0.359)
Dark-lighted	–	–	–	–	–	–	0.307 (0.094)	–
Dark-unlighted	–	–	–	–	–	−1.005 (0.518)	–	–

Days of week								
Weekend	–	–	–	–	–	–	–	−0.442 (0.172)
Crash characteristics								
Presence of passenger (base: no passenger)								
One passenger	–	0.876 (0.137)	–	–	–	0.300 (0.091)	–	–
Two passenger	−0.342 (0.122)	0.421 (0.144)	–	–	–	–	–	–
More than two passengers	−0.342 (0.122)	–	–	–	–	–	−0.265 (0.086)	–
Trajectory of vehicle's motions (base: other movement)								
Going straight	–	–	–	0.505 (0.063)	0.312 (0.089)	0.505 (0.063)	0.505 (0.063)	–

[a] Standard errors are presented in parenthesis.

in newer vehicles (vehicle age < 6 years); this can be attributed to the absence of advanced safety features in older vehicles.

Roadway design attributes: In terms of roadway design attributes, the estimates indicate that crashes on gravel road surface tend to be less severe compared to crashes on paved and unpaved surfaces for head-on collision. On gravel road surfaces, drivers are compelled to drive cautiously at a slower speed contributing to a reduction in the severity of crash outcomes. It is very interesting to note that the presence of signal decreases the injury propensity for rear-ender collision, and increases the injury propensity for both angular collisions (near- and far-angular). Injury propensity reductions are observed for the presence of pedestrian control (for rear-ended and short side-angular), roundabouts (for near- and short side-angular), stop sign (for short side-angular) and yield sign (for rear-ender and short side-angular).

The results for speed zones indicate that the drivers are likely to sustain severe injuries for crashes occurring in zones with medium and higher speed limits highlighting that the probability of sustaining severe injuries increases with the increasing speed limits – a surrogate for vehicle speed at the time of crash. Among the type of intersection variables, T-intersection leads to higher injury propensity for rear-ender collisions and lower injury propensity for far-sideswipe collisions. Five or more legged intersection reflects reduced injury risk propensity for rear-ended collision. The reduction is also observed for non-intersection location in rear-ended and short side angular collision propensities.

Environmental factors: In the category of environmental factors; time of day, weather condition and lighting condition have significant influence in moderating the driver injury severity across different collision types. With respect to the time of day, higher severity levels are associated with head-on collision during morning peak period. As expected, the injury severity for drivers is higher during late night. This is particularly so for rear-ender, short side-angular and head-on collision. The injury risk propensities of near-sideswipe and far-angular collision reflect higher severities for rainy/snowy/foggy weather. This may be due to unfavorable driving conditions resulting from the reduced visibility during adverse weather conditions. For high wind condition, rear-end collision propensities (rear-ender and rear-ended) indicate lower likelihood of severe injuries. The parameter characterizing the effect of weekend suggests lower injury severity level for head-on collision. The result is quite interesting and the reasons for the effect are not very clear. It is possibly a manifestation of unobserved information that is not considered in our analysis and warrants additional investigation in the future.

Crash characteristics: Presence of passenger and trajectory of vehicle's motions are the crash characteristics that are found to affect driver injury severity. A higher injury risk propensity is observed for the presence of one passenger in the vehicle for the rear-ended and far-angular collision. However, the result associated with two passengers has a more uneven effect across different collision types indicating lower and higher likelihood of severe injury in the effect of rear-ender and rear-ended propensities, respectively. But presence of more than two passengers indicates lower likelihood of severe injury for rear-ender and short side-angular collision. Overall, the drivers with the presence of more passengers are less likely to be severely injured presumably a reflection of more responsible driving behavior in the presence of passengers (the same effect is observed in Eluru et al., 2010). Finally, the coefficients corresponding to the vehicle movement reveal that straight vehicle movement of the driver increases the injury risk propensity compared to other turning movements for far-sideswipe, near-, far- and short side-angular collisions. The result is expected because the drivers are likely to be traveling at a higher speed while traveling straight.

Table 5
Elasticity effects for collision type component.

Variables	Rear-ender	Rear-ended	Near-sideswipe	Far-sideswipe	Near-angular	Far-angular	Short-side angular	Head-on
Driver characteristics								
Driver age (base: age 25–64)								
Age less than 25	40.052	−36.538	0.811	0.888	0.653	0.667	0.731	0.124
Age above 65+	−2.373	−36.715	−0.452	−53.674	25.389	26.211	−7.004	2.607
Driver gender (base: male)								
Female	−9.030	36.468	−8.155	−7.794	13.883	−9.601	−9.606	−29.141
Restraint system use (base: seat belt used)								
Seat belt not used	36.388	−5.688	−6.026	−6.029	−4.373	−4.443	−4.466	−6.411
Vehicle characteristics								
Vehicle type (base: sedan)								
Utility	25.514	−6.913	−8.598	−8.698	−4.668	−4.571	−4.588	18.495
Panel van	63.287	−9.834	−10.378	−10.395	−7.638	−7.763	−7.805	−11.037
Vehicle age (base: vehicle age less than 6)								
Vehicle age 6–10	11.740	−6.122	−21.807	11.030	−19.612	−5.395	12.873	9.223
Vehicle age 11 and above	9.808	−6.815	−28.154	−35.500	−4.457	−4.685	9.346	9.712
Roadway design attributes								
Type of road surface (base: paved)								
Gravel	−15.978	−14.490	−21.988	−22.344	−7.502	−6.705	−6.714	137.276
Traffic control device (base: none traffic control and other control device)								
Signal	17.724	15.759	−43.360	−43.319	8.339	7.798	7.796	−98.954
Pedestrian control	114.393	111.707	−46.430	−46.886	−37.835	−38.792	−39.048	−46.820
Roundabout	−13.374	17.187	−13.015	−12.459	37.679	−14.206	−14.138	−10.784
Stop sign	−54.852	−70.433	24.128	−92.304	157.516	−7.493	−7.133	−90.032
Yield sign	0.846	−28.630	−95.608	−95.711	110.263	−11.532	−11.315	−78.830
Speed zone (base: low speed ≤50 km/h)								
Medium speed (60–90 km/h)	39.964	39.521	−18.763	−18.811	−12.982	−13.048	−13.139	−19.165
High speed (≥100 km/h)	27.923	28.489	22.347	33.058	−19.898	−19.922	−20.066	47.918
Type of intersection (base: cross intersection)								
T intersection	−6.852	14.867	5.415	24.146	−2.274	−27.260	−14.652	96.935
Five and more legged intersection	18.932	51.882	−25.865	136.337	−17.955	−18.533	−18.744	−27.104
Non-intersection	76.774	74.714	103.364	96.499	−24.854	−83.150	−82.944	220.728
Environmental factors								
Time of day (base: morning peak, off peak and late evening)								
Late night	4.898	4.921	4.290	4.265	6.109	−27.427	6.525	2.695
Surface condition (base: dry)								
Wet	−9.591	−8.540	−13.453	−13.741	−4.305	−3.780	−3.765	80.177
Lighting condition (base: daylight)								
Dusk/dawn	1.660	1.652	2.153	−42.425	0.984	1.043	1.067	2.577
Dark-lighted	−28.355	−14.310	9.538	9.565	6.756	6.863	6.928	9.638
Dark-unlighted	−10.604	−9.515	−14.745	−15.028	−4.866	−4.309	−4.304	89.813
Days of week (base: weekdays)								
Weekend	−2.347	−26.895	−4.597	39.239	0.062	0.337	0.357	41.697

5. Elasticity effects and validation analysis

The parameter estimates of Tables 3 and 4 do not provide the magnitude of the effects of exogenous variables on the probability of involving in a specific type of collision or sustaining a specific injury severity category for drivers, respectively. For this purpose, we compute the aggregate level "elasticity effects" for all independent variables (see Eluru and Bhat (2007) for a discussion on the methodology for computing elasticities). The effects are computed for both the collision type and injury severity components and are presented in Tables 5 and 6, respectively. However, to conserve on space, we present the elasticity effects only for the highest injury severity level (serious/fatal injury severity category) across all collision types.

The following observations can be made based on the results presented in Tables 5 and 6. *First*, the most significant variables in terms of collision type are: crashes at non-intersection location, crashes on gravel roads, presence of pedestrian control, driving a panel van, driver age less than 25, medium speed limit zone and not wearing seat-belt. *Second*, the most significant variables in terms of increase in serious/fatal injury for drivers are crashes in high speed limit zone and driver age 65 and above. In terms of serious/fatal injury reduction, the important factors are driving a station wagon, presence of roundabout and presence of pedestrian control. *Third*, the impacts, in magnitude, are substantially different in injury severity for several variables (driver age 65+, non-local driver, high speed limit road and collision during late-night) across different collision types. The effects are also different in direction (sign) for presence of signal and collision at T intersection. These differences clearly highlight that each collision type has a fundamentally distinct injury severity profile underscoring the importance of examining the effect of various exogenous variables on driver injury severity outcome by different collision types.

In an effort to further assess the performance of the joint model, a validation experiment is also carried out. For testing the predictive performance of the models, 50 data samples, of about 5000 records each, are randomly generated from the hold out validation sample consisting of 25,769 records. For

Table 6
Elasticity effects for serious/fatal injury severity category.

Variables	Rear-ender	Rear-ended	Near-sideswipe	Far-sideswipe	Near-angular	Far-angular	Short-side angular	Head-on
Driver characteristics								
Driver age (base: age 25–64)								
Age less than 25	–	−35.757	–	–	–	–	−16.532	–
Age above 65+	48.599	–	128.737	–	54.896	48.825	59.236	–
Driver gender (base: male)								
Female	72.080	65.571	64.029	71.814	63.669	72.314	68.000	58.566
Restraint system use (base: seat belt used)								
Seat belt not used	–	–	–	–	63.399	–	–	–
Non-local driver	–	27.496	–	81.777	24.752	–	–	–
Vehicle characteristics								
Vehicle type (base: sedan)								
Station wagon	−38.650	−20.505	–	−71.215	−36.310	−20.673	−37.388	−16.409
Utility	–	–	–	–	–	–	−47.841	–
Panel van	–	–	–	–	–	–	–	−48.777
Vehicle age (base: vehicle age less than 6)								
Vehicle age 6–10	20.437	–	18.066	–	–	20.804	19.551	–
Vehicle age 11 and above	27.753	28.177	24.822	28.376	25.345	–	26.529	–
Roadway design attributes								
Type of road surface (base: paved)								
Gravel	–	–	–	–	–	–	–	−35.466
Traffic control device (base: none traffic control and other control device)								
Signal	−32.734	–	–	–	20.103	59.027	–	–
Pedestrian control	–	−60.601	–	–	–	–	−59.060	–
Roundabout	–	–	–	–	−68.876	–	−69.494	–
Stop sign	–	–	–	–	–	–	−31.320	–
Yield sign	−61.729	–	–	–	–	–	−25.443	–
Speed zone (base: low speed ≤50 km/h)								
Medium speed (60–90 km/h)	–	–	–	–	–	30.204	34.311	–
High speed (≥100 km/h)	–	–	80.545	89.029	89.666	174.732	153.362	64.888
Type of intersection (base: cross intersection)								
T intersection	23.907	–	–	−79.906	–	–	–	–
Five and more legged intersection	−62.106	–	–	–	–	–	–	–
Nonintersection	–	−22.965	–	–	–	–	−20.703	–
Environmental factors								
Time of day (base: morning peak, off peak and Late evening)								
Morning peak	–	–	–	–	–	–	–	55.937
Late night	177.750	–	–	–	–	–	53.022	106.305
Weather condition (base: clear)								
Rain/snow/fog/smoke/dust	–	–	70.554	–	–	87.743	–	–
High wind	−53.838	−53.830	–	–	–	–	–	–
Lighting condition (base: daylight)								
Dusk/dawn	–	–	–	–	–	–	27.162	50.379
Dark-lighted	–	–	–	–	–	–	28.979	–
Dark-unlighted	–	–	–	–	–	−62.290	–	–
Crash characteristics								
Presence of passenger (base: no passenger)								
One passenger	–	75.635	–	–	–	28.133	–	–
Two passenger	–	43.947	–	–	–	–	–	–
More than two passengers	–	–	–	–	–	–	−22.103	
Days of week (base: weekdays)								
Weekend	–	–	–	–	–	–	–	−30.567
Trajectory of vehicle's motions (base: other movement)								
Going straight	–	–	–	46.105	26.585	47.230	40.339	–

these samples, we present the average measures of predictive log-likelihood and BIC values along with the 95% level confidence band. The average predictive log-likelihood measure for the copula model and independent model are −13277.24 [(−13326.17) to (−13228.30)] and −13280.37 [(−13329.306) to (−13231.438)], respectively. The BIC values for the copula model and independent model are 27714.26 [27615.130–27813.394] and 27720.13 [27621.79–27818.47], respectively, further highlighting the enhanced performance of the copula model.

6. Conclusions

The focus of this paper is to jointly model the collision type and injury severity outcome of drivers involved in a two vehicle collisions using a copula-based joint multinomial logit-ordered logit modeling framework. The current study contributes to the literature on driver injury severity in three ways. The *first contribution* of our research is to define a vehicle level collision type variable using a combination of collision type and the initial point of contact. The *second contribution* of our study is to develop a closed form copula based framework to accommodate the impact of observed and unobserved effects on collision type and injury severity while generating collision type as a vehicle level variable. Finally, our *third contribution* is to formulate the copula model by incorporating parameterization of dependency profile in an unordered and ordered joint structure. The proposed model is estimated using driver injury severity data for two vehicle crashes from the state of Victoria, Australia employing a comprehensive set of exogenous variables – driver characteristics, vehicle characteristics, roadway design attributes, environmental factors and crash characteristics.

The empirical analysis involves estimation of models by using six different copula structures: (1) Gaussian, (2) FGM, (3) Clayton, (4) Gumbel, (5) Frank and (6) Joe. The most suitable copula model is obtained for a combination model of Frank–Clayton copulas (Frank copula structure for rear-ender and head-on collision and Clayton dependency structure with the remaining collision type). Further, the comparison between copula and the independent models confirms the importance of accommodating dependence between collision type and injury severity outcome in the analysis of driver injury severity. The model estimation results presented in the current paper suggest that the impact of exogenous variables vary (for some variables) in magnitude as well as in sign across collision types. The variables in moderating the effect of different collision types also reveal varying effects.

In our research, to further understand the impact of various exogenous factors, elasticity effects are estimated for both the collision type and injury severity components. The elasticity effects clearly highlight that each collision type has a fundamentally distinct injury severity profile underscoring the importance of examining the effect of various exogenous variables on driver injury severity outcome by different collision types. In summary, the findings of this paper provide a more complete picture of injury severity profile associated with different collision type, thus target based countermeasures could be devised to address the entire profile of collision mechanism.

Acknowledgements

The corresponding author would like to acknowledge financial support from Natural Sciences and Engineering Research Council (NSERC) Discovery grants program and Fonds de recherche du Québec - Nature et technologies (FQRNT) Team Grant Program for undertaking the research. The authors would also like to acknowledge useful feedback from four anonymous reviewers.

References

Abay, K.A., Paleti, R., Bhat, C.R., 2013. The joint analysis of injury severity of drivers in two-vehicle crashes accommodating seat belt use endogeneity. Transportation Research Part B: Methodological 50, 74–89.

Abdel-Aty, M.A., Abdelwahab, H.T., 2003. Configuration analysis of two-vehicle rear-end crashes. In: Transportation Research Record: Journal of the Transportation Research Board, No. 1840. Transportation Research Board of the National Academies, Washington, DC, pp. 140–147.

Alexander, J., Barham, P., Black, I., 2002. Factors influencing the probability of an incident at a junction: results from an interactive driving simulator. Accident Analysis and Prevention 34 (6), 779–792.

Bédard, M., Guyatt, G.H., Stones, M.J., Hirdes, J.P., 2002. The independent contribution of driver, crash, and vehicle characteristics to driver fatalities. Accident Analysis and Prevention 34 (6), 717–727.

Bhat, C.R., 1997. Covariance heterogeneity in nested logit models: econometric structure and application to intercity travel. Transportation Research Part B: Methodological 31 (1), 11–21.

Bhat, C.R., 2011. The Maximum Approximate Composite Marginal Likelihood (MACML) estimation of multinomial probit-based unordered response choice models. Transportation Research Part B: Methodological 45 (7), 923–939.

Bhat, C.R., Eluru, N., 2009. A copula-based approach to accommodate residential self-selection effects in travel behavior modeling. Transportation Research Part B: Methodological 43 (7), 749–765.

Chamberlain, G., 1980. Analysis of covariance with qualitative data. The Review of Economic Studies 47 (1), 225–238.

Chiou, Y.C., Hwang, C.C., Chang, C.C., Fu, C., 2013. Modeling two-vehicle crash severity by a bivariate generalized ordered probit approach. Accident Analysis & Prevention 51, 175–184.

Chipman, M.L., 2004. Side impact crashes – factors affecting incidence and severity: review of the literature. Traffic Injury Prevention 5 (1), 67–75.

Conroy, C., Tominaga, G.T., Erwin, S., Pacyna, S., Velky, T., Kennedy, F., Sise, M., Coimbra, R., 2008. The influence of vehicle damage on injury severity of drivers in head-on motor vehicle crashes. Accident Analysis and Prevention 40 (4), 1589–1594.

Das, A., Abdel-Aty, M.A., 2011. A combined frequency-severity approach for the analysis of rear-end crashes on urban arterials. Safety Science 49 (8–9), 1156–1163.

Eluru, N., 2013. Evaluating alternate discrete choice frameworks for modeling ordinal discrete variables. Accident Analysis & Prevention 55, 1–11.

Eluru, N., Bhat, C.R., 2007. A joint econometric analysis of seat belt use and crash-related injury severity. Accident Analysis & Prevention 39 (5), 1037–1049.

Eluru, N., Bhat, C.R., Hensher, D.A., 2008. A mixed generalized ordered response model for examining pedestrian and bicyclist injury severity level in traffic crashes. Accident Analysis & Prevention 40 (3), 1033–1054.

Eluru, N., Paleti, R., Pendyala, R.M., Bhat, C.R., 2010. Modeling injury severity of multiple occupants of vehicles: copula-based multivariate approach. In: Transportation Research Record: Journal of the Transportation Research Board, No. 2165. Transportation Research Board of the National Academies, Washington, DC, pp. 1–11.

Fredette, M., Mambu, L.S., Chouinard, A., Bellavance, F., 2008. Safety impacts due to the incompatibility of suvs, minivans, and pickup trucks in two-vehicle collisions. Accident Analysis and Prevention 40 (6), 1987–1995.

Gårder, P., 2006. Segment characteristics and severity of head-on crashes on two-lane rural highways in Maine. Accident Analysis & Prevention 38 (4), 652–661.

Islam, S., Mannering, F., 2006. Driver aging and its effect on male and female single-vehicle accident injuries: some additional evidence. Journal of Safety Research 37 (3), 267–276.

Jin, Y., Wang, X., Chen, X., 2010. Right-angle crash injury severity analysis using ordered probability models. In: 2010 International Conference on Intelligent Computation Technology and Automation, ICICTA 2010, pp. 206–209.

Khattak, A.J., 2001. Injury severity in multivehicle rear-end crashes. In: Transportation Research Record: Journal of the Transportation Research Board, No. 1746. Transportation Research Board of the National Academies, Washington, DC, pp. 59–68.

Kim, J.-K., Ulfarsson, G.F., Kim, S., Shankar, V.N., 2013. Driver-injury severity in single-vehicle crashes in california: a mixed logit analysis of heterogeneity due to age and gender. Accident Analysis & Prevention 50, 1073–1081.

Lee, L.-F., 1983. Generalized econometric models with selectivity. Econometrica 51 (2), 507–512.

Loo, B.P.Y., Tsui, K.L., 2007. Factors affecting the likelihood of reporting road crashes resulting in medical treatment to the police. Injury Prevention 13 (3), 186–189.

Mooradian, J., Ivan, J.N., Ravishanker, N., Hu, S., 2013. Analysis of driver and passenger crash injury severity using partial proportional odds models. Accident Analysis & Prevention 58, 53–58.

Nordhoff, L.S., 2005. Motor Vehicle Collision Injuries: Biomechanics, Diagnosis, and Management. Jones and Bartlett Learning.

Portoghese, A., Spissu, E., Bhat, C.R., Eluru, N., Meloni, I., 2011. A copula-based joint model of commute mode choice and number of non-work stops during the commute. International Journal of Transport Economics 38 (3), 337–362.

Quinn, C., 2007. The health-economic applications of copulas: Methods in applied econometric research. In: Health, Econometrics and Data Group (HEDG) Working Paper 07/22. Department of Economics, University of York.

Rana, T., Sikder, S., Pinjari, A., 2010. Copula-based method for addressing endogeneity in models of severity of traffic crash injuries. In: Transportation Research Record: Journal of the Transportation Research Board, No. 2147. Transportation Research Board of the National Academies, Washington, DC, pp. 75–87.

Risbey, T., Cregan, M., de Silva, H., 2010. Social Cost of Road Crashes. Australasian Transport Research Forum, Canberra.

Savolainen, P.T., Mannering, F.L., Lord, D., Quddus, M.A., 2011. The statistical analysis of highway crash-injury severities: a review and assessment of methodological alternatives. Accident Analysis & Prevention 43 (5), 1666–1676.

Schiff, M.A., Cummings, P., 2004. Comparison of reporting of seat belt use by police and crash investigators: variation in agreement by injury severity. Accident Analysis & Prevention 36 (6), 961–965.

Schneider, W., Savolainen, P., Zimmerman, K., 2009. Driver injury severity resulting from single-vehicle crashes along horizontal curves on rural two-lane highways. In: Transportation Research Record: Journal of the Transportation Research Board, No. 2102. Transportation Research Board of the National Academies, Washington, DC, pp. 85–92.

Sener, I.N., Eluru, N., Bhat, C.R., 2010. On jointly analyzing the physical activity participation levels of individuals in a family unit using a multivariate copula framework. Journal of Choice Modelling 3 (3), 1–38.

Sklar, A., 1973. Random variables, joint distribution functions, and copulas. Kybernetika 9, 449–460.

Terza, J.V., 1985. Ordinal probit: a generalization. Communications in Statistics: Theory and Methods 14 (1), 1–11.
Trivedi, P.K., Zimmer, D.M., 2007. Copula modeling: an introduction for practitioners. Foundations and Trends in Econometrics 1 (1), 1–110.
Tsui, K.L., So, F.L., Sze, N.N., Wong, S.C., Leung, T.F., 2009. Misclassification of injury severity among road casualties in police reports. Accident Analysis & Prevention 41 (1), 84–89.
Ulfarsson, G.F., Mannering, F.L., 2004. Differences in male and female injury severities in sport-utility vehicle, minivan, pickup and passenger car accidents. Accident Analysis & Prevention 36 (2), 135–147.
Washington, S., Karlaftis, M.G., Mannering, F.L., 2003. Statistical and Econometric Methods for Transportation Data Analysis. Chapman & Hall/CRC, Boca Raton.
WHO, 2013. Global Status Report on Road Safety 2013. World Health Organization.
Williams, A.F., Shabanova, V.I., 2003. Responsibility of drivers, by age and gender, for motor-vehicle crash deaths. Journal of Safety Research 34 (5), 527–531.
Yan, X., Radwan, E., Abdel-Aty, M., 2005. Characteristics of rear-end accidents at signalized intersections using multiple logistic regression model. Accident Analysis & Prevention 37 (6), 983–995.
Yasmin, S., Eluru, N., 2013. Evaluating alternate discrete outcome frameworks for modeling crash injury severity. Accident Analysis & Prevention 59, 506–521.
Ye, X., Pendyala, R.M., Al-Rukaibi, F.S., Konduri, K., 2008. A joint model of accident type and severity for two-vehicle crashes. In: Presented at 87th Annual Meeting of the Transportation Research Board. Washington, DC.
Zhang, C., Ivan, J.N., 2005. Effects of geometric characteristics on head-on crash incidence on two-lane roads in Connecticut. In: Transportation Research Record: Journal of the Transportation Research Board, No. 1908. Transportation Research Board of the National Academies, Washington, DC, pp. 159–164.
Zuxuan, D., Ivan, J.N., Gårder, P., 2006. Analysis of factors affecting the severity of head-on crashes two-lane rural highways in Connecticut. In: Transportation Research Record: Journal of the Transportation Research Board, No. 1953. Transportation Research Board of the National Academies, Washington, DC, pp. 137–146.