Evaluating Temporal Variability of Exogenous Variable Impacts over 25 Years: An Application of Scaled Generalized Ordered Logit Model for Driver Injury Severity

Robert Marcoux
Undergraduate Student
Department of Civil, Environmental and Construction Engineering
University of Central Florida
Tel: 321-276-7580, Fax: 407-823-3315
Email: BobbyMrcx@Knights.ucf.edu

Shamsunnahar Yasmin*
Postdoctoral Associate
Department of Civil, Environmental and Construction Engineering
University of Central Florida
Tel: 407-823-4815, Fax: 407-823-3315
Email: shamsunnahar.yasmin@ucf.edu

Naveen Eluru
Associate Professor
Department of Civil, Environmental and Construction Engineering
University of Central Florida
Tel: 407-823-4815, Fax: 407-823-3315
Email: naveen.eluru@ucf.edu

Moshiur Rahman
Doctoral Student
Department of Civil, Environmental and Construction Engineering
University of Central Florida
Tel: 321-276-7580, Fax: 407-823-3315
Email: moshiur@knights.ucf.edu

---

* Corresponding author

ABSTRACT

The current study undertakes a unique research effort to quantify the impact of various exogenous factors on crash severity over time. Specifically, we examine if over time, the impact of exogenous variables has changed and if so what is the magnitude of the change. The research contributes to driver injury severity analysis both methodologically and empirically by proposing a framework that addresses the challenges associated with pooled (or pseudo-panel) data. For our analysis, we draw data from the General Estimates System (GES) over a span of twenty-five years. The data is compiled for driver injury severity in single or two vehicle crashes from 1989 through 2014 in 5-year increments (1989, 1994, 1999, 2004, 2009 and 2014). The alternative econometric frameworks considered for the analysis include ordered logit, generalized ordered logit, scaled generalized ordered logit and mixed generalized ordered logit models. A host of comparison metrics are computed to evaluate the performance of these alternative models in examining the pooled data. The model development exercise is conducted with a host of exogenous variables including driver characteristics, vehicle characteristics, roadway attributes, environmental factors, crash characteristics and temporal attributes. The model estimation results are further augmented by performing a detailed policy scenario analysis, probability profile representations and elasticity effects for different driving and situational conditions across different years.

Keywords: Driver injury severity; Temporal instability; Pseudo-panel; Generalized ordered logit; Mixed Model; Scaled Model

INTRODUCTION

The negative consequences of road traffic crashes have a significant impact on the emotional and financial well-being of the society. In the United States in 2015, roadway crashes were responsible for over 35,000 fatalities, while about 2.35 million individuals were injured or disabled. Financially, road traffic crashes cost the US nearly 230 billion annually. Given the magnitude of consequences of road traffic crashes, it is not surprising that road safety is a well-researched field. Earlier research has examined factors affecting crash occurrence using crash frequency models (see Yasmin and Eluru, 2016 for a review) and crash consequence in the event of a crash using crash severity models (see Yasmin and Eluru, 2013 for a review). The current study contributes to road safety literature by focusing on crash severity models. Earlier literature on severity modeling has identified several factors that significantly influence severity of vehicle occupants involved in traffic crashes including vehicle occupant age, restraint system use, driving under the influence of alcohol or drugs, vehicle age and type (such as sedan, van or pickup truck) of the vehicles involved, and collision type (such as head-on, rear-end and angle).

To be sure, road safety in the US has improved over the years. The number of traffic fatalities have reduced from about 47,000 in 1965 to about 35,000 in 2015 (NHTSA, 2016). The fatality rate per 100 million vehicle miles travelled has dropped from 5.3 to 1.12 during the same period. The improvement in traffic crash associated fatalities can chiefly be attributed to: (1) design and enforcement of several policies such as mandatory seat belt use, vehicle regulations requiring airbags, child rear facing seats, (2) advances in vehicle technology to improve occupant safety and (3) concerted effort dedicated to educational awareness campaigns for different driver age groups to encourage safe driving behavior. However, in recent years (since 2000) instances of increase in the number of traffic fatalities compared to the previous year have occurred multiple times. Thus, in spite of the significant progress made over the years, there is further scope for improving road safety.

The current study undertakes a unique research effort to quantify the impact of various exogenous factors on crash severity over time (see Mannering, 2018 for discussion on temporal instability). Specifically, we examine if over time, the impact of exogenous variables has changed and if so what is the magnitude of the change. For example, the injury severity of a vehicle occupant in a crash that occurred in 1995 was a function of the vehicle safety features available at the time. An observationally identical crash occurring in 2015 is likely to result in either the same level of injury severity or lower. This is an example of how improvement in vehicle technology has affected injury severity. In our study, we systematically attempt to identify exogenous variables that offer time-varying effects and quantify the change in their impact. For this purpose, we draw data from the General Estimates System (GES), over a span of twenty-five years. The data is compiled for driver injury severity in single or two passenger vehicle involved crashes from 1989 through 2014 in 5-year increments (1989, 1994, 1999, 2004, 2009 and 2014). In our analysis, injury severity is classified in four levels as follows: no injury, possible injury, non-incapacitating injury, and incapacitating/fatal injury.

The data compiled is a pooled dataset obtained from stitching together 6 cross-sectional datasets providing us with a pseudo-panel data. Such data pooling of different observations across multiple years offers unique methodological challenges. The modeling methodology should recognize the differences across multiple time points adequately since the outcome process for the observations in a year might be influenced by various observed and unobserved attributes (Anowar et al., 2016; Train, 2009). For illustrative purposes, consider the possibility that airbags were made mandatory in vehicles only in 2000. The data compiled after this time period would possibly experience lower injury severity for crashes relative to the years prior due to the installation of airbags. This is an example of an observed attribute affecting either one or more cross-sections of data in the overall pooled dataset. On the other hand, consider the possibility of a cultural phenomenon that encourages good driving behavior – such as wearing a seatbelt – happened between 1995 and 2000. As a result, more drivers wear seatbelts after 2000, thus potentially reducing injury severity consequences for cross-sections compiled later. However, parsing the true impact of the cultural phenomenon in the model is quite challenging and usually remains hidden or unobserved for a long time. This is an example of how an unobserved effect specific to one-time period or multiple time periods can affect severity outcomes. In our study, we implement modeling approaches that simultaneously accommodate for the influence of observed and unobserved attributes on driver injury severity across multiple time points.

Given the inherent ordering of the data, we adopt a generalized ordered logit (GOL) model kernel structure for severity analysis. The GOL framework relaxes the restrictive assumption of the traditional ordered outcome (Ordered logit/probit) models (monotonic effect of exogenous variables) while simultaneously recognizing the inherent ordering of the injury severity variable - information that unordered model alternatives fail to consider (see Eluru et al., 2008; Eluru and Yasmin, 2015; Yasmin and Eluru, 2013). Further, to incorporate the effect of observed and unobserved temporal effects, we specifically consider two versions of the GOL model – the mixed GOL model and the scaled GOL model. The two variants differ in the way they incorporate the influence of unobserved attributes within the outcome process. The mixed GOL model captures unobserved heterogeneity by allowing the variable effects (including the constant) influencing injury severity to be distributed across the observations, while the scaled model accommodates for common unobserved heterogeneity by allowing the variance of the unobserved component to vary by time period. The two approaches also vary in how they accommodate for heteroscedasticity across observations. We estimate both models and employ data fit comparison metrics to determine the appropriate model structure. The model specification is undertaken to quantify how the impact of exogenous variables has altered over the 25-year period on driver injury severity. The model development exercise is conducted with a host of exogenous variables including driver characteristics, vehicle characteristics, roadway attributes, environmental factors, crash characteristics and temporal attributes. In summary, the current research effort contributes empirically by identifying if the impact of any exogenous factors has varied over time, and if so, quantifying the change in magnitude. Methodologically, the research contributes to driver injury severity analysis by proposing a framework that addresses the challenges associated with pooled (or pseudo-panel) data.

The remainder of the paper is organized as follows. Section 2 provides a brief literature review. Section 3 briefly outlines the econometric frameworks that are considered. Section 4 presents data preparation steps for the modeling exercise. In section 5, we discuss the model results. Section 6 presents a summary of policy analysis. Finally, Section 6 provides the summary and concludes our findings.

LITERATURE REVIEW

Several research efforts have examined driver injury severity in safety literature. It is beyond the scope of the paper to provide a detailed review (see Savolainen et al., 2011; Yasmin and Eluru, 2013 for a detailed review). Based on earlier literature, the most important factors influencing injury severity by attribute group are presented below with emphasis on factors contributing to or reducing injury severity outcome in the event of a crash.

Driver Characteristics: Driver characteristics including old driver indicator, female driver, alcohol impairment, aggressive driving and not wearing seat-belt are found to increase probability of serious crashes (Valent et al., 2002; Yasmin et al., 2014). Young driver, male driver and wearing seat belt are predominantly found in safety literature to be associated with lower crash injury severity outcomes (Donmez and Liu, 2015; Eluru et al., 2010; Fredette et al., 2008).

Vehicle Characteristics: Among different vehicle characteristics, vehicle type and vehicle age are the two most common attributes considered in evaluating injury severity. Higher vehicle age and car are likely to result in higher injury severity outcomes (Kockelman and Kweon, 2002; Xie et al., 2009). On the other hand, heavy passenger vehicles, such as pickup, utility vehicle and light truck are found to decrease higher severity outcomes for the occupants of these vehicles (Srinivasan, 2002; Wu et al., 2016).

Roadway Attributes: Non-intersection related crashes and higher speed limits are roadway attributes which increase likelihood of injury severity outcomes (Al-Ghamdi, 2002; Yasmin et al., 2015). On the other hand, crash at a traffic signal, increase in traffic flow and protected left turn phase variables are found to reduce severe crash likelihood (Christoforou et al., 2010; Quddus et al., 2009; Rifaat et al., 2011; Wang and Abdel-Aty, 2008).

Environmental Factors: Rainy weather, peak hour, daylight and snowy road surface decreases the likelihood of severe crashes (Anderson and Hernandez, 2017; Chen et al., 2016; Huang et al., 2008) while night-time crash, darkness with no road lighting condition result in higher crash severity outcomes (Huang et al., 2008; Khattak, 2001).

Crash Characteristics: Collision type is often considered as an important crash characteristic in evaluating injury severity outcomes (Yasmin et al., 2014). Among different collision types, rollover, single vehicle crash, head-on crash, angular crashes are associated with higher injury severity outcome (Eluru and Bhat, 2007; Krull et al., 2000), while rear-end crash and same direction sideswipe crash result in lower crash severity outcome (Obeng, 2008; Ouyang et al., 2002).

While earlier research has identified the impact of several attributes there has been limited research conducted to explore temporal evolution of exogenous variables on injury severity. Two studies conducted temporal stability analysis of injury severity estimates. Behnood and Mannering (2015) conducted a detailed study of parameter stability over time using data from Chicago for the years 2004 through 2012. The authors employed a mixed multinomial logit model and estimated separate models by year. To compare model parameters across years they computed marginal effects and presented graphical comparisons of these effects. In another study Dabbour (2017) examined stability of injury severity models for data of North Carolina from 2007 through 2013 within a logistic regression framework. The study estimated separate models for each year and compared the mode estimates as well as the odds ratios of the parameters. While these two studies offer insights on temporal parameter stability, their approaches to parameter comparison do not yield conclusive evidence on the change in parameters. To elaborate, while marginal effects were compared, there is no consideration for the confidence band of these marginal effects. Without these confidence bands, it is not possible to conclude if the differences in marginal effects is a true difference or a random occurrence. In addition, the parameter differences was considered separately and information on how the parameters vary in time cannot be generated based on the separate model for each year. Further, the year range considered in these studies is between 7 to 9 years. While the timeframe is significantly large, it is still not adequate to capture very long term trends in parameter evolution.

The current study makes multiple contributions by building on the work of Behnood and Mannering (2015) while explicitly addressing some of the salient aspects highlighted in Mannering (2018). First, we expand the time span of parameter stability analysis by compiling data from 6 time points over a 25-year span from 1989 through 2014. Second, instead of estimating separate models for each year, we stitch the data together to obtain a pseudo-panel of crash records. The pseudo-panel data is examined using two generalized versions of the GOL model that can control for year specific effects as well as unobserved factors: (1) scaled GOL and (2) mixed GOL. In the current study context, scaled GOL model is specified by allowing the variance of the unobserved component to vary across different years, while the mixed GOL model is specified by allowing the unobserved component to vary across different observations. The scaled GOL has a closed form expression and does not require simulation for model estimation in accommodating for temporal unobserved heterogeneity. On the other hand, the mixed GOL model requires adoption of maximum simulated likelihood approach for estimation. In estimating pooled models, variables representing "time elapsed from 1989" which is the time difference between the analysis year (1989, 1994, 1999, 2004, 2009, and 2014) from the base year (1989) considered in the current study context. Moreover, interaction of exogenous variables with the time elapsed variable are utilized to control for time varying variable effects. As a result, it would be possible to identify how the parameters vary in time based on a single pooled model rather than making inferences from separate year-specific models. In addition to enhancing variable inference, the approach also makes it easier to evaluate marginal effects across multiple years. The pooled approach for model estimation will provide conclusive evidence of changes to the model parameters when interacted with time indicator variable. If the interaction term is insignificant, we can conclude that there is no temporal variability. Finally, we employ parametric forms of parameter estimate differences across multiple points allowing us to extrapolate the parameters into the future.

ECONOMETRIC FRAMEWORK

In this section, we provide a brief description of the methodology of all the models considered for examining driver injury severity in our research. For the ease of presentation, we describe the general mathematical structure first and then identify the different modeling structures for various models in the subsequent discussion. We will first introduce the traditional ordered logit (OL) model, then discuss about the generalized ordered logit (GOL) model, scaled generalized ordered logit (SGOL) model, and finally present the mixed version of the generalized ordered logit (MGOL) model.

Let us assume that q $(q = 1,2, \dots, Q)$ be the index for driver and k $(k = 1,2, \dots, K)$ be the index for injury severity categories. In this empirical study, k takes the value of 'no injury' $(k = 1)$, 'possible injury' $(k = 2)$, 'non-incapacitating injury $(k = 3)$, and 'incapacitating injury/fatal injury' $(k = 4)$. In the traditional ordered outcome model, the discrete injury severity levels (y_q) are assumed to be associated with an underlying continuous latent variable (y_q^*). This latent variable is typically specified as the following linear function:

$$y_q^* = \alpha' x_q + \varepsilon_q , \qquad y_q = k \qquad if\ \psi_{k-1} < y_q^* < \psi_k \tag{1}$$

where y_q^* is mapped to the injury severity level y_q by the ψ thresholds ($\psi_0 = -\infty$ and $\psi_k = \infty$) in the usual ordered-outcome fashion. x_q is a column vector of attributes (not including a constant) that influences the propensity associated with injury severity outcome. α' is a corresponding column vector of coefficients and ε_q is an idiosyncratic random error term assumed to be identically and independently standard logistic distributed across driver q. Given these relationships across the different parameters, the resulting probability expressions for individual q and alternative k for the OL take the following form:

$$P_q(k) = \Lambda\big(\psi_k - \alpha' x_q\big) - \Lambda\big(\psi_{k-1} - \alpha' x_q\big) \quad (2)$$

where $\Lambda(.)$represents the standard logistic cumulative distribution function (cdf).

GOL is a flexible form of the traditional OL model that relaxes the restriction of constant threshold across population. The GOL model represents the threshold parameters as a linear function of exogenous variables (Eluru et al., 2008; Srinivasan, 2002). In order to ensure the ordering of observed discrete injury severity levels $\big(-\infty < \psi_{q,1} < \psi_{q,2} < \cdots \ldots \ldots < \psi_{q,k-1} < +\infty\big)$, we employ the following parametric form followed by Eluru et al. (2008):

$$\psi_{q,k} = \psi_{q,k-1} + exp(\gamma_{qk} + \delta'_{qk} Z_{qk}) \quad (3)$$

where, Z_{qk} is a set of explanatory variables associated with the k^{th} threshold (excluding a constant), δ'_{qk} is a vector of parameters to be estimated and γ_{qk} is a parameter associated with injury severity level k. The remaining structure and probability expressions are similar to the OL model. For identification reasons, we need to restrict one of the δ'_k vectors to zero.

For both OL and GOL models, the probability expression of Equation 2, is derived by assuming that the variance in propensity over different injury severity levels across different years is unity. However, we can introduce a *scale parameter (λ)*, which would scale the coefficients to reflect the variance of the unobserved portion of the utility for each time point. Thus, the probability expression can be written as:

$$P_q(k) = \Lambda\left[\frac{\big(\psi_k - \alpha' x_q\big)}{\lambda}\right] - \Lambda\left[\frac{\big(\psi_{k-1} - \alpha' x_q\big)}{\lambda}\right] \quad (4)$$

where λ is the scale parameter of interest and is parameterized as $\exp(\sigma \mathrm{t_i})$ and t_i is a time elapsed variable which is the time difference between the analysis year (1989, 1994, 1999, 2004, 2009, and 2014) from the base year (1989) considered. Thus the variable takes the form of an ordered variable ranging from 0 (for 1989) to 25 (for 2014). This yields the SGOL model. If the σ parameters are not significantly different from 0, the expression in equation (4) collapses to the expression in equation (2) yielding either the OL or GOL model depending on the threshold characterization. The reader would note that the scale parameter also relaxes the homoscedasticity assumption implicit within the OL and GOL formulations.

The mixed GOL accommodates unobserved heterogeneity in the effect of exogenous variables on injury severity levels in both the latent injury severity propensity function and the threshold functions (Eluru et al., 2008; Srinivasan, 2002). The equation system for MGOL model can be expressed as:

$$y_q^* = (\alpha' + \beta') x_q + \varepsilon_q \quad (5)$$

$$\psi_{q,k} = \psi_{q,k-1} + exp[(\delta'_{qk} + \theta'_{qk}) Z_{qk}] \quad (6)$$

We assume that β' and θ'_{qk} are independent realizations from normal distribution for this study. The proposed approach takes the form of a random coefficients GOL model thus allowing us to capture the influence of exogenous variables and year specific error correlation through elements of x_q and Z_{qk}. This approach is analogous to splitting the error term (ε_q) into multiple error components (analogous to error components mixed logit model). The parameters to be estimated in the MGOL model are the mean and covariance matrix of the distributions of β' and θ'_{qk}. In this study, we use the Halton sequence (200 Halton draws) to evaluate the multidimensional integrals (see Eluru et al., 2008 for a similar estimation process). In our analysis, x_q vector includes the year of the data collection allowing us to estimate observed and unobserved variations with respect to time.

DATA

Data Source and Data Description

The data for the current study is sourced from the "General Estimates System (GES)" database for the years 1989, 1994, 1999, 2004, 2009, and 2014. The GES database is a nationally representative sample of road crashes collected and compiled from about 60 jurisdictions across the United States. The data is obtained from the U. S. Department of Transportation, National Highway Traffic Safety Administration's National Center for Statistics and Analysis (ftp://ftp.nhtsa.dot.gov/GES/). The data includes information on reports compiled by police officers for crashes involving at least one motor vehicle traveling on a roadway and resulting in property damage, injury or death to the road users. For the six years, the crash database has a record of 286,490 crashes involving 493,249 vehicles and 789,576 individuals. A five-point ordinal scale is used in the database to represent the injury severity of individuals involved in these crashes[1]: 1) No injury; 2) Possible injury; 3) Non-incapacitating injury; 4) Incapacitating injury and 5) Fatal injury. Further, the dataset compiles information on a multitude of factors (driver characteristics, vehicle characteristics, roadway attributes, environmental factors, and crash characteristics) representing the crash situations and events. A number of crash-related factors are extracted from this database in order to explore the variables that might influence the driver injury severity.

Data Assembly and Data Preparation

The focus of this study is on injury severity of drivers of passenger vehicles (passenger car, sport utility vehicle, pickup or van). Thus, the following criteria were employed for sample formation: (1) The crashes that involve only non-commercial (private) passenger vehicle drivers are selected (to avoid the potential systematic differences between commercial and non-commercial driver groups). (2) The passenger vehicle crashes that involve another passenger vehicle or a fixed object are examined. The crashes that involve more than two vehicles are excluded from the analysis. The final dataset of non-commercial driver of passenger vehicles from all years, after removing records with missing information for essential attributes consisted of about 251,701 records – with 31,012 records for the year 1989; 42,858 records for the year 1994; 45,959 records for the year 1999; 46,889 records for the year 2004; 31,408 records for the year 2009 and 53,575 records for

[1] Injury severity levels are defined and considered based on NHTSA (2015). Specifically, in the pooled dataset, "no injury and non apparent injury" are defined as no injury, "non-incapacitating evident injury and suspected minor injury" are defined as non-incapacitating injury; "incapacitating injury and suspected serious injury" are defined as incapacitating injury. The definition of possible and fatal injury are same across all years.

the year 2014. From this dataset, a sample of 2000 records from each year is sampled out for the purpose of analysis. Thus, the final estimation sample has 12,000 data records. Table 1 presents the sample share for different severity levels in the estimation sample across different years considered. From the table, we can see that the injury severity shares remain reasonably stable over the years, however for 2004 and 2009, the incapacitating and fatal crash shares are higher relative to other years. In this final sample, the percentage of fatal crashes sustained by drivers is extremely small (0.5%). Therefore, both the fatal and incapacitating injury categories are merged together to ensure a representative share for each alternative crash level. In the final estimation sample, the distributions of driver injury severities are: no injury 70.7%, possible injury 13.3%, non-incapacitating injury 10.0% and incapacitating/fatal injury 6.1%.

MODEL ESTIMATION RESULTS

Variable Considered

In road crash reporting system, the definition of independent variables and the reporting approach may vary across different years. The Analytical User's Manual of GES data has detailed information on the label of each variables reported in the GES database from 1988 to 2014 (NHTSA, 2015). From the report it is quite evident that there are variations in variable documentation across different years. However, the variables should be uniform for all years in a pooled dataset for estimating a single pooled model. Therefore, in our current study, to estimate models using pooled data, we prepared the datasets such that all years have exactly the same set of independent variables with same number of levels in each category. To reiterate, we have considered only the common attributes present in different years and did not consider the variables which were omitted or added in the crash database across years. In our study, we considered a host of exogenous variables and divided those into six broad categories: Driver characteristics (including driver gender, driver age, alcohol consumption and restraint system use), Vehicle characteristics (including vehicle type and vehicle age), Roadway attributes (including roadway class, road location and traffic control device), Environmental factors (including time of day, day of week, lighting condition and road surface condition), and Crash characteristics (including collision object and manners of collision). Finally, in terms of Temporal variables, we introduced a variable called "time elapsed from 1989" which is the time difference between the most recent years (1994, 1999, 2004, 2009, and 2014) from the base year (1989) considered in the current study context. Both linear and square effects of the time elapsed were tested. Moreover, interaction of exogenous variables with the time elapsed variable (linear and square) were utilized to control for time varying variable effects. As a result, it would be possible to apply the developed models for future year scenarios. Table 2 offers a summary of the sample characteristics of the categorical exogenous factors in the estimation dataset. The only continuous variable considered in our study is vehicle age which has a mean of 7.45 years for different years considered. The final specification was based on a systematic process of removing statistically insignificant variables (90% confidence level) and combining variables when their effects were not significantly different. The specification process was also guided by prior research, intuitiveness and parsimony considerations.

Overall Measures of Fit

The empirical analysis involves the estimation of four models: (1) the ordered logit (OL) model, (2) the generalized ordered logit (GOL) model, (3) the scaled generalized ordered logit (SGOL) model, and (4) the mixed generalized ordered logit (MGOL) model. Prior to discussing the estimation results, we compare the performance of these models in this section. The log-likelihood values at convergence for the various frameworks are as follows: (1) OL (with 31 parameters) is -10463.112; (2) GOL (with 47 parameters) is -10384.608, (3) SGOL (with 49 parameters) is -10380.072, and (4) MGOL (with 51 parameters) is -10380.468. The corresponding value for the "constant only" model is -10963.882. To undertake the comparison, we employ a two-step process. In the first step, we use the likelihood-ratio (LR) test to determine the two superior models among all the models estimated (The ordered models (OL, GOL, SGOL and MGOL) are nested version of each other and hence LR test is appropriate). Subsequently, we compare the selected two models by using several information criteria to determine the best fitted models.

The LR test statistic is computed as $2[LL_U - LL_R]$, where LL_U and LL_R are the log-likelihood of the unrestricted and the restricted models, respectively. The computed value of the LR test is compared with the $\aleph^2$ value for the corresponding degrees of freedom (*dof*). The resulting LR test values for the comparison of OL/GOL, OL/SGOL, OL/MGOL, GOL/SGOL and GOL/MGOL models are 157.008 (16 *dof*), 166.080 (18 *dof*), 165.288 (20 *dof*), 9.072 (2 *dof*) and 8.280 (4 *dof*), respectively. The LR test values indicate that MGOL and SGOL outperform both OL and GOL models indicating that MGOL and SGOL offer superior fit compared to both OL and GOL models. To further evaluate the performance of SGOL and MGOL models, we employ different information criterion including: 1) Bayesian Information Criterion (BIC), 2) Akaike Information Criterion (AIC) and 3) Akaike Information Criterion corrected (AICc). These measures can be computed as:

$$\begin{aligned} BIC &= -2LL + K\, ln(Q) \\ AIC &= 2K - 2ln(LL) \\ AIC_c &= 2K - 2ln(LL) + \frac{2\,K(K+1)}{(Q - K - 1)} \end{aligned} \tag{7}$$

where LL is the log-likelihood value at convergence, K is the number of parameters, and Q is the number of observations. The BIC (AIC; AICc) values for the final specifications of the SGOL and MGOL models are 21218.807 (20858.144; 20858.554) and 21243.865 (20862.936; 20863.380), respectively. The comparison exercise clearly highlights the superiority of the SGOL model, in current study context, in terms of data fit compared to all the other models.

Estimation Results

In presenting the effects of exogenous variables in the model specification, we will restrict ourselves to the discussion of the SGOL model. Table 3 presents the estimation results. In SGOL model, when the threshold parameter is positive (negative), the result implies that the threshold is bound to increase (decrease); the actual effect on the probability is quite non-linear and can only be judged in conjunction with the influence of the variable on propensity and other thresholds. In the following sections, the estimation results are discussed by variable groups.

Temporal Variables

With respect to time elapsed variables, both linear and square effects of the time elapsed variables are found to be significant determinants of driver injury severity outcomes. The linear specification of time elapsed variable is significant in the propensity and in both thresholds. The effect in the propensity for the time elapsed variable indicates that driver injury severity in recent times are likely to be severe. But the impact of time elapsed variable on the thresholds indicate that the temporal effect on severe crash categories are crash and driver-specific. The square specification of time elapsed variable implies negative effect on driver injury severity with an overall less likelihood of incapacitating/fatal crash outcomes (as indicated by second threshold). Therefore, we can argue that the net effect of time elapsed on driver injury severity can be identified based on combination of both linear and square specifications of the variable with their impact in propensities and thresholds. To have a better understanding of time elapsed variable on driver injury severity outcome, we plot the injury severity probabilities for different years. The plot is presented in Figure 1. From Figure 1, we can see that no injury probability has increased substantially over time while other injury severity outcomes have decreased over time. The overall improvement in injury severity outcomes may be attributed to improvement in technology, infrastructure, enforcement and education. It is also evident from the figure that the higher injury severity outcomes have shown a slight increase in recent years, which might be attributed to distracted driving – a topic of concern highlighted within the safety community..

Driver Characteristics

The demographics of driver involved in the collision has significant influence on crash severity. The result related to driver gender indicates that compared to the male drivers, the latent injury propensity is higher for female drivers. The result is perhaps indicating that females are less capable of bearing physical and mental trauma compared to males (Chen and Chen, 2011; Sivak et al., 2010). The estimates associated with driver age, suggest a reduction in the likelihood of severe injuries for the young drivers (age<25) compared to middle-aged drivers (age 25 to 64). The lower probability of severe injury among young adults may reflect the higher physiological strength of young drivers in withstanding crash impacts (Castro et al., 2013; Xie et al., 2012). Further, a negative impact of the interaction term between the young driver and time elapsed (square specification) is observed in our analysis. Consistent with earlier studies (Bédard et al., 2002; Kim et al., 2013), we also find that old drivers (age≥65) involved in crashes are less likely to evade serious injury relative to other adult individuals. Further, the negative sign of threshold of old driver demarcating the possible and non-incapacitating injury indicates a higher likelihood of non-incapacitating and incapacitating/fatal injuries for this group of drivers. As expected, drivers under the influence of alcohol are likely to have a higher injury risk propensity compared to the sober drivers. Negative sign of the threshold demarcating the possible and non-incapacitating shows a higher likelihood of non-incapacitating and incapacitating/fatal injury. The impact of use of restraint system does not affect propensity, however, a positive impact of the interaction term between unrestrained and time elapsed is found significant in the threshold demarcating the possible injury and non-incapacitating injury. On the other hand, the unrestrained variable with the interaction of time elapsed with square specification is found significant with negative signs in both thresholds. The actual effect is non-linear and can only be judged in conjunction with the influence of the variable with both the linear and square specifications of time elapsed effect.

Vehicle Characteristics

With respect to vehicle characteristics, vehicle type and vehicle age are found to be significant determinates of driver injury severity. The driver's vehicle type indicators reveal that the drivers of pickup, van, light truck and utility vehicles have lower injury risk propensity relative to car, perhaps due to the larger weight of these vehicles (see Eluru et al., 2010; Xie et al., 2009 for similar results). At the same time, the negative values of the first thresholds of pickup reflects an increase in non-incapacitating injury probability. On the other hand, the positive sign in the threshold specific to van implies lower likelihood of non-incapacitating and other severe injury categories. The contrasting effect of propensity and threshold for pickup implies that the effect of pickup on the non-incapacitating and incapacitating/fatal injury categories is crash and driver-specific; for some contexts, the minor injury probability can increase with a concomitant decrease in the serious/fatal injury probability, while for other contexts the reverse can hold. This highlights the advantage of a GOL framework that allows for flexible exogenous variable impacts. The vehicle age estimate demonstrates that drivers are more likely to be severely injured if they drive older vehicles while involved in a road crash. Several previous studies have also demonstrated such impacts of older vehicles on the outcome of driver injury severity (Islam and Mannering, 2006; Kim et al., 2013).

Roadway Attributes

With respect to the roadway class, the model estimates show that the likelihood of severe injury increases for the driver of passenger vehicles when the crash occurs on an interstate highway. From the result of the interaction of interstate highway with time elapsed variable (square specification), we observe that with time, injury severity for crashes occurring on interstate highway is decreasing. The result is perhaps indicating better roadway design over years. Among other roadway attributes, road location and type of traffic control device have significant influence of driver injury severity profiles. Specifically, intersection related, or driveway access or other roadway location related crashes are less likely to result in severe injuries to the drivers in the event of a crash relative to non-intersection and intersection locations. The intersection related variable is also found to have significant effects in both thresholds with positive signs. The net implication is that intersection related crashes have a lower probability of resulting in fatal crashes. Crashes in the presence of stop-sign seem to decrease injury severity likelihood relative to other traffic control systems.

Environmental Factors

Several environmental factors considered in our model are found to be significant determinants of driver injury severity in the final model specification. The likelihood of severe injury is lower during evening peak period compared to other time of day. The lower crash severity outcome of evening peak may be attributed to higher traffic volume and slow driving speeds during this period of the day. The estimation results also reveal that drivers are likely to evade severe injury during weekends relative to weekday. The lighting condition effect shows a higher probability of severe injury crashes during dark-unlighted conditions, perhaps due to reduced visibility and higher reaction time during darkness. The indicator variable also reveals an overall exacerbating effects on severe crash outcomes as indicated by the negative sign of the first threshold. The surface condition effects indicate that if collisions occur on a wet or other road surface (relative to dry surface conditions), the drivers are more likely to evade injury; possibly because of cautious driving during adverse weather condition.

Crash Characteristics
Collision with large object (building, concrete traffic barrier, wall, tree, bridge, snow bunk) has a positive effect on the propensity of injury severity, while also demonstrating a higher likelihood of non-incapacitating injury (related to collision with small object and moving vehicle). The result is in line with several previous studies (Holdridge et al., 2005; Yamamoto and Shankar, 2004). With regards to manner of collision, the results in Table 3 related to sideswipe-same direction collisions are likely to result in lower injury risk propensities relative to other collision types. On the other hand, the impacts of the indicator variable on the first threshold is negative, which implies that the effects of sideswipe-same direction collision on different injury categories are crash and driver-specific. Overall, the results suggest an increased probability of possible injury category.

The result associated with a head-on collision reflects an increased likelihood of severe crash outcome. The pre-impact speed vectors of motor vehicles are directed in opposing directions during a head-on collision, resulting in greater dissipation of kinetic energy and heavier deformation of motor vehicle bodies, resulting in higher risk of injury (Gårder, 2006; Tay and Rifaat, 2007). The effect of head-on collision type is also negative in the threshold demarcating non-incapacitating and incapacitating/fatal injury which indicates higher likelihood of incapacitating/fatal injury outcomes. From the result of the interaction of head-on collision with time elapsed variable (linear specification), we observe that with time, injury severity resulting from head-on crashes is decreasing, which can be attributed to better safety protection designs of vehicles over time[2]. The positive sign of propensity associated with angular collision reflects higher likelihood of severe injury, which can be attributed to the greater force of impact in an angular collision (Tay and Rifaat, 2007).

Scale Parameter
As indicated earlier, in SGOL model specification, we introduce a scale parameter to reflect the variance of the unobserved portion for each time point. In current study context, we use time elapsed (linear specification) variable as scale parameter and from Table 3, we can see that the variable effect is significant. The scale parameter indicates significant variation in the unobserved factors across the years. Specifically, there is an increased influence of unobserved factors on injury severity with time relative to the first time period. The higher scale value can be attributed to increased influence of variables not considered in our models such as vehicle technology and roadway infrastructure improvements, and education and enforcement campaigns. The result supports our hypothesis that there is significant variability in the variance of the unobserved component of driver injury severity propensity across different time periods.

POLICY ANALYSIS

Model Illustration

The exogenous variable coefficients do not directly provide the magnitude of impacts of variables on the probability of different injury severity levels. Moreover, the impacts of coefficients of the SGOL framework might not be readily interpretable due to the interactions between propensity, threshold and scale parameters. Hence, to provide a better understanding of the impacts of exogenous factors, we compute disaggregate level changes in driver injury severity levels. Specifically, we focus on a hypothetical scenario (based on the variables found significant in the final model specifications).

[2] In severity analysis, collision type is often endogenous to severity outcomes (see Rana et al., 2010 and Yasmin et al., 2014). However, in our current study, it is beyond the scope to examine for such endogeneity.

Let us consider a driver involved in an intersection related crash, during evening peak period on a wet road surface condition. For this hypothetical condition, we generate probability profiles for a young driver (aged less than 25) for different years by changing other crash attributes. The probability profile would allow us to understand the variation in injury severity profile over time and across different situations. In generating the probability profile, we consider the following conditions:

1) <u>Driver condition 1:</u> Young driver, driving a pickup which is 7-year-old, driver is not wearing seat belt and driving while impaired by alcohol
2) <u>Driver condition 2:</u> Young driver, driving a pickup which is 2-year-old, driver is not wearing seat belt and driving while impaired by alcohol
3) <u>Driver condition 3:</u> Young driver, driving a pickup which is 2-year-old and the driver is not wearing seat belt, while not impaired by alcohol
4) <u>Driver condition 4:</u> Young driver, driving a pickup which is 2-year-old and the driver is wearing seat belt while not impaired by alcohol
5) <u>Driver condition 5:</u> Young driver, driving an utility passenger vehicle which is 2-year-old and the driver is wearing seat belt while not impaired by alcohol
6) <u>Driver condition 6:</u> Young driver, driving a car which is 2-year-old and the driver is wearing seat belt while not impaired by alcohol

For these driver conditions, the probability plots are generated for all six years considered in the analysis (1989, 1994, 1999, 2004, 2009 and 2014) along with a future year (2019) and these plots are presented in Figure 2. The reader would note that the probability plots provided are only a sample of the various illustrations that can be generated based on the independent variables in the models.

In Figure 2, predicted probabilities (Z-axis) identified based on SGOL model estimation results are depicted by 3-dimensional (3D) bar plots as a function of the injury severity levels (X-axis), the year (Y-axis) and probability (Z-axis). Overall, from the plots we can observe that probability of serious injury severity outcomes of young driver for alcohol impairment and not wearing seat belt conditions are greater in recent years. A similar trend is also observed for old vehicle indicator variable as well. As is expected, among three unsafe conditions explored (old vehicle, alcohol impaired and not wearing seat belt), alcohol impairment has greater impact on serious crash outcome for young driver – evident by the drop-in incapacitating/fatal bars from driving condition 2 to 3. For newer vehicle, the severity profiles are almost similar across different passenger vehicles (pickup, utility and car). However, it is interesting to note that driving newer vehicles are likely to have deteriorating impacts on severe crash outcomes for the interim years considered (1999, 2004, 2009), and in recent years the positive impacts of newer vehicles appear to be restored. The generated probability plots clearly show that over the years the safety trends have changed. The development of such injury severity profiles could be helpful for the policy makers to identify problematic conditions and identifying countermeasures for improving driver safety.

Elasticity Effects

The model illustration section provides a representation of how overall injury severity profiles vary over time. However, to identify the impact of various exogenous variables, we compute the aggregate level "elasticity effects" for independent variables (see Eluru and Bhat, (2007) for a discussion on the methodology for computing elasticities). We present the elasticity effects for a set of variables – Female, Under the influence of alcohol, Unrestrained, Pickup, On interstate highway, Dark-unlighted, Collision with large object and Head on collision. The elasticity effects are presented by severity categories for different years considered in the analysis along with the mean elasticity effects and confidence interval of effects at 90% level. The results in the table can be interpreted as the percentage change (increase for positive sign and decrease for negative sign) in the probability of the crash severity categories due to the change in that specific exogenous variable.

The following observations can be made based on the elasticity effects of the variables presented in Table 4. First, among the variables considered for elasticity computation, the most significant variables in terms of increase in incapacitating/fatal injury are head on collision, collision with large object and driving under the influence of alcohol. These variables are also the most significant contributors to incapacitating/fatal injury in the most recent year (2014) considered. Second, from base year to the most recent year considered, the effect of head on crash and crashes on interstate highway on incapacitating/fatal injury have decreased significantly. On the other hand, the effect of driving while unrestrained has increased significantly with regards to incapacitating/fatal injury. Third, the effects of under the influence of alcohol and unrestrained driving on serious injury severity categories (non-incapacitating/incapacitating/fatal injury) have seen significant increase over the years considered. Fourth, there are substantial differences in the elasticity effects across different years considered. Finally, the confidence interval generated for the mean elasticity highlights the presence of relatively narrow bands for all variables across various severity categories. The elasticity analysis conducted provides an illustration on how the proposed model can be employed to identify the most crucial issues in improving overall driver safety and in understanding the critical factors contributing to driver injury severity over time.

CONCLUSION

The current study undertook a unique research effort to quantify the impact of various exogenous factors on crash severity over time. Specifically, we examined if over time, the impact of exogenous variables has changed and if so what is the magnitude of the change. In our study, we systematically identify exogenous variables that offer time-varying effects and quantify the change in their impact. For this purpose, we drew data from the General Estimates System (GES) database, over a span of twenty-five years. The data was compiled for driver injury severity in single or two vehicle (passenger vehicle only) crashes from 1989 through 2014 in 5-year increments (1989, 1994, 1999, 2004, 2009 and 2014). In our analysis, injury severity was classified in four levels as follows: no injury, possible injury, non-incapacitating injury, and incapacitating/fatal injury. The data compiled is a pooled dataset obtained from stitching together 6 cross-sectional datasets providing us with a pseudo-panel data. Such data pooling of different observations across multiple years offers unique methodological challenges. The modeling methodology should recognize the differences across multiple time points adequately since the outcome process for the observations in a year might be influenced by various observed and unobserved attributes. Towards that end, in our study, we implemented modeling approaches that simultaneously accommodate for the influence of observed and unobserved attributes on driver injury severity across multiple time points.

Given the inherent ordering of the data, we estimated models by using generalized ordered logit (GOL), scaled generalized ordered logit (SGOL), and mixed generalized ordered logit (MGOL) model structures along with a traditional ordered logit model. These model development exercises were conducted with a host of exogenous variables including driver characteristics, vehicle characteristics, roadway attributes, environmental factors, crash characteristics and temporal attributes and the focus of this study was injury severity of drivers of passenger vehicles (passenger car, sport utility vehicle, pickup or van). The comparison exercise based on log-likelihood ratio test and different information criteria clearly highlighted the superiority of the estimated SGOL model in terms of data fit compared to all the other models. Further, effects of the elapsed time from the base year (1989) of data collection (and their interaction with other observed variables) were found significant which supported our hypothesis that there is variation in variable impacts on driver injury severity across the years. From the estimation results, we found that though the injury severity outcome had seen significant decrease over time, there were a slight increase in the categories of severe injury severities in recent years.

In our research, to further understand the impact of various exogenous factors, we conducted a detailed policy analysis and 3-dimensional representation of injury severity profile considering different hypothetical driving and situational conditions across different years. The generated probability plots clearly indicated that over the years safety trends have changed. The development of such injury severity profiles could be helpful for the policy makers to identify problematic conditions and identifying countermeasures for improving driver safety. Further, we presented and discussed elasticities of different variables across different years to demonstrate the implications of the estimated model. From the elasticity effects, we found that the effect of head on crash and crashes on interstate highway on incapacitating/fatal injury have decreased significantly over time. On the other hand, the effects of under the influence of alcohol and unrestrained driving on serious injury severity categories (non-incapacitating/incapacitating/fatal injury) have seen significant increase over different years considered. These results warrant stricter enforcement to prevent driving under the influence of alcohol and without seat-belt in improving the overall driver safety.

In our study, we estimated mixed GOL and scaled GOL model separately, while it is also possible to specify these models in a single framework by incorporating both unobserved effects by observations and by time period. In future, it might be interesting to examine crash injury severity by employing a GOL model that accounts for scale variation within a mixed GOL model. Moreover, it might be useful to specify the temporal effect as a combination of parametric and categorical design to capture further year-specific insights. In future, it might be interesting to examine crash injury severity by employing a scaled-mixed GOL model as a combination of parametric and categorical design of the temporal effect. Finally, while our proposed approach focused on severity propensity, it might be interesting, in a future effort, to evaluate for the potential influence of collision type endogeneity on severity propensity over time (see Rana et al., 2010 and Yasmin et al., 2014).

REFERENCES

Al-Ghamdi, A.S., 2002. Using logistic regression to estimate the influence of accident factors on accident severity. *Accident Analysis and Prevention* 34(6), 729-741.

Anderson, J., Hernandez, S., 2017. Roadway classifications and the accident injury severities of heavy-vehicle drivers. *Analytic Methods in Accident Research* 15, 17-28.

Anowar, S., Eluru, N., Miranda-Moreno, L.F., 2016. Analysis of vehicle ownership evolution in Montreal, Canada using pseudo panel analysis. *Transportation* 43(3), 531-548.

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.

Behnood, A., Mannering, F.L., 2015. The temporal stability of factors affecting driver-injury severities in single-vehicle crashes: Some empirical evidence. *Analytic Methods in Accident Research* 8, 7-32.
Castro, M., Paleti, R., Bhat, C.R., 2013. A spatial generalized ordered response model to examine highway crash injury severity. *Accident Analysis and Prevention* 52, 188-203.
Chen, C., Zhang, G., Liu, X.C., Ci, Y., Huang, H., Ma, J., Chen, Y., Guan, H., 2016. Driver injury severity outcome analysis in rural interstate highway crashes: a two-level Bayesian logistic regression interpretation. *Accident Analysis and Prevention* 97, 69-78.
Chen, F., Chen, S., 2011. Injury severities of truck drivers in single-and multi-vehicle accidents on rural highways. *Accident Analysis and Prevention* 43(5), 1677-1688.
Christoforou, Z., Cohen, S., Karlaftis, M.G., 2010. Vehicle occupant injury severity on highways: an empirical investigation. *Accident Analysis and Prevention* 42(6), 1606-1620.
Dabbour, E., 2017. Investigating temporal trends in the explanatory variables related to the severity of drivers' injuries in single-vehicle collisions. *Journal of Traffic and Transportation Engineering (English Edition)* 4(1), 71-79.
Donmez, B., Liu, Z., 2015. Associations of distraction involvement and age with driver injury severities. *Journal of safety research* 52, 23-28.
Eluru, N., Bhat, C.R., 2007. A joint econometric analysis of seat belt use and crash-related injury severity. *Accident Analysis and 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 and Prevention* 40(3), 1033-1054.
Eluru, N., Paleti, R., Pendyala, R., Bhat, C., 2010. Modeling injury severity of multiple occupants of vehicles: Copula-based multivariate approach. *Transportation Research Record: Journal of the Transportation Research Board* (2165), 1-11.
Eluru, N., Yasmin, S., 2015. A note on generalized ordered outcome models. *Analytic methods in accident research* 8, 1-6.
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 and Prevention* 38(4), 652-661.
Holdridge, J.M., Shankar, V.N., Ulfarsson, G.F., 2005. The crash severity impacts of fixed roadside objects. *Journal of Safety Research* 36(2), 139-147.
Huang, H., Chin, H.C., Haque, M.M., 2008. Severity of driver injury and vehicle damage in traffic crashes at intersections: A Bayesian hierarchical analysis. *Accident Analysis and Prevention* 40(1), 45-54.
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.
Khattak, A., 2001. Injury severity in multivehicle rear-end crashes. *Transportation Research Record: Journal of the Transportation Research Board* (1746), 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 and Prevention* 50, 1073-1081.
Kockelman, K.M., Kweon, Y.-J., 2002. Driver injury severity: an application of ordered probit models. *Accident Analysis and Prevention* 34 (3), 313-321.

Krull, K., Khattak, A., Council, F., 2000. Injury effects of rollovers and events sequence in single-vehicle crashes. *Transportation Research Record: Journal of the Transportation Research Board* (1717), 46-54.
Mannering, F., 2018. Temporal instability and the analysis of highway accident data. *Analytic Methods in Accident Research*, 17, 1-13.
NHTSA, 2016. *2015 Motor vehicle crashes: overview, traffic safety facts: research notes,* National highway Traffic Safety Adminstration (NHTSA), U. S. Department of Transportation.
NHTSA, 2015. *National Automotive Sampling System (NASS) General Estimates System (GES) Analytical User's Manual 1988–2014*, National Highway Traffic Safety Adminstration (NHTSA), U. S. Department of Transportation.
Obeng, K., 2008. Injury severity, vehicle safety features, and intersection crashes. *Traffic injury prevention* 9(3), 268-276.
Ouyang, Y., Shankar, V., Yamamoto, T., 2002. Modeling the simultaneity in injury causation in multivehicle collisions. *Transportation Research Record: Journal of the Transportation Research Board* (1784), 143-152.
Quddus, M.A., Wang, C., Ison, S.G., 2009. Road traffic congestion and crash severity: econometric analysis using ordered response models. *Journal of Transportation Engineering* 136 (5), 424-435.
Rana, T., S. Sikder, and A. Pinjari, 2010. Copula-based method for addressing endogeneity in models of severity of traffic crash injuries. *Transportation Research Record* (2147), 75-87.
Rifaat, S.M., Tay, R., De Barros, A., 2011. Effect of street pattern on the severity of crashes involving vulnerable road users. *Accident Analysis and Prevention* 43(1), 276-283.
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 and Prevention* 43(5), 1666-1676.
Sivak, M., Schoettle, B., Rupp, J., 2010. Survival in fatal road crashes: body mass index, gender, and safety belt use. *Traffic injury prevention* 11(1), 66-68.
Srinivasan, K., 2002. Injury severity analysis with variable and correlated thresholds: ordered mixed logit formulation. *Transportation Research Record: Journal of the Transportation Research Board* (1784), 132-141.
Tay, R., Rifaat, S.M., 2007. Factors contributing to the severity of intersection crashes. *Journal of Advanced Transportation* 41(3), 245-265.
Train, K.E., 2009. *Discrete choice methods with simulation*. Cambridge university press.
Valent, F., Schiava, F., Savonitto, C., Gallo, T., Brusaferro, S., Barbone, F., 2002. Risk factors for fatal road traffic accidents in Udine, Italy. *Accident Analysis and Prevention* 34(1), 71-84.
Wang, X., Abdel-Aty, M., 2008. Analysis of left-turn crash injury severity by conflicting pattern using partial proportional odds models. *Accident Analysis and Prevention* 40(5), 1674-1682.
Wu, Q., Zhang, G., Ci, Y., Wu, L., Tarefder, R.A., Alcántara, A.D., 2016. Exploratory multinomial logit model–based driver injury severity analyses for teenage and adult drivers in intersection-related crashes. *Traffic injury prevention* 17(4), 413-422.
Xie, Y., Zhang, Y., Liang, F., 2009. Crash injury severity analysis using Bayesian ordered probit models. *Journal of Transportation Engineering* 135(1), 18-25.
Xie, Y., Zhao, K., Huynh, N., 2012. Analysis of driver injury severity in rural single-vehicle crashes. *Accident Analysis and Prevention* 47, 36-44.

Yamamoto, T., Shankar, V.N., 2004. Bivariate ordered-response probit model of driver's and passenger's injury severities in collisions with fixed objects. *Accident Analysis and Prevention* 36(5), 869-876.
Yasmin, S., Eluru, N., 2013. Evaluating alternate discrete outcome frameworks for modeling crash injury severity. *Accident Analysis and Prevention* 59, 506-521.
Yasmin, S., Eluru, N., 2016. Latent segmentation based count models: analysis of bicycle safety in Montreal and Toronto. *Accident Analysis and Prevention* 95, 157-171.
Yasmin, S., Eluru, N., Pinjari, A.R., 2015. Pooling data from fatality analysis reporting system (FARS) and generalized estimates system (GES) to explore the continuum of injury severity spectrum. *Accident Analysis and Prevention* 84, 112-127.
Yasmin, S., Eluru, N., Pinjari, A.R., Tay, R., 2014. Examining driver injury severity in two vehicle crashes – A copula based approach. *Accident Analysis and Prevention* 66, 120-135.

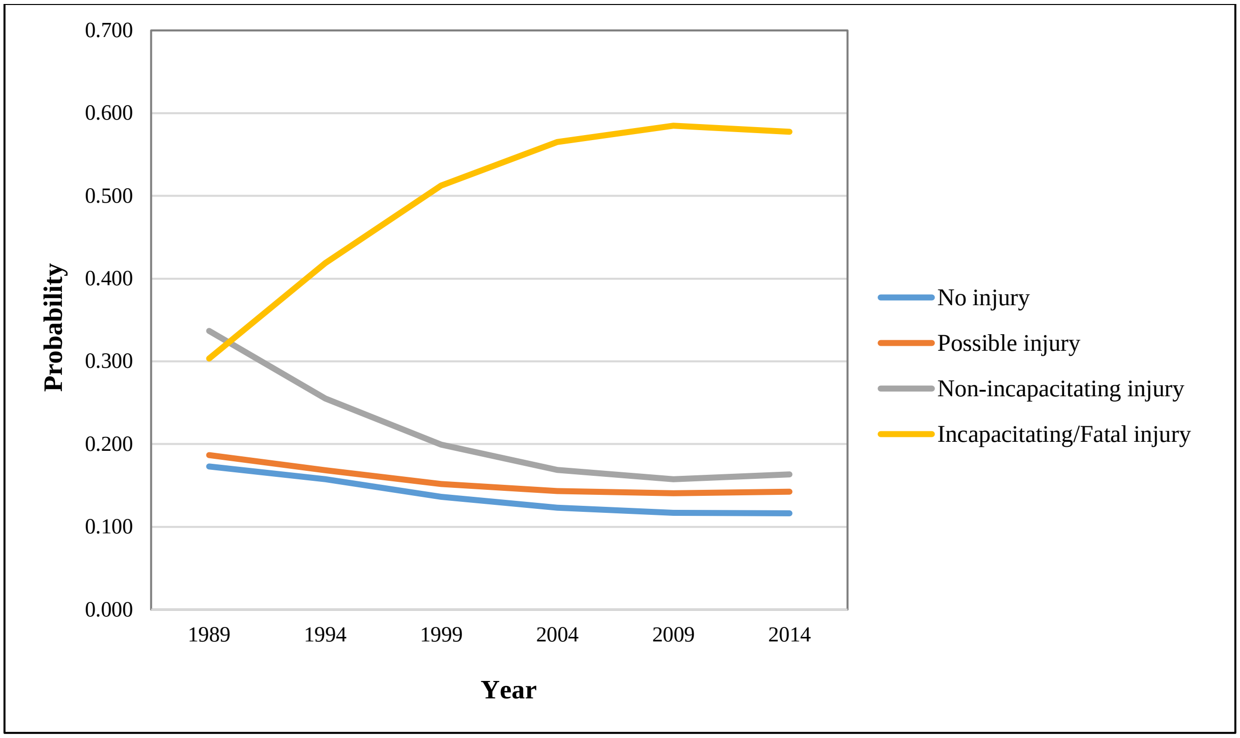
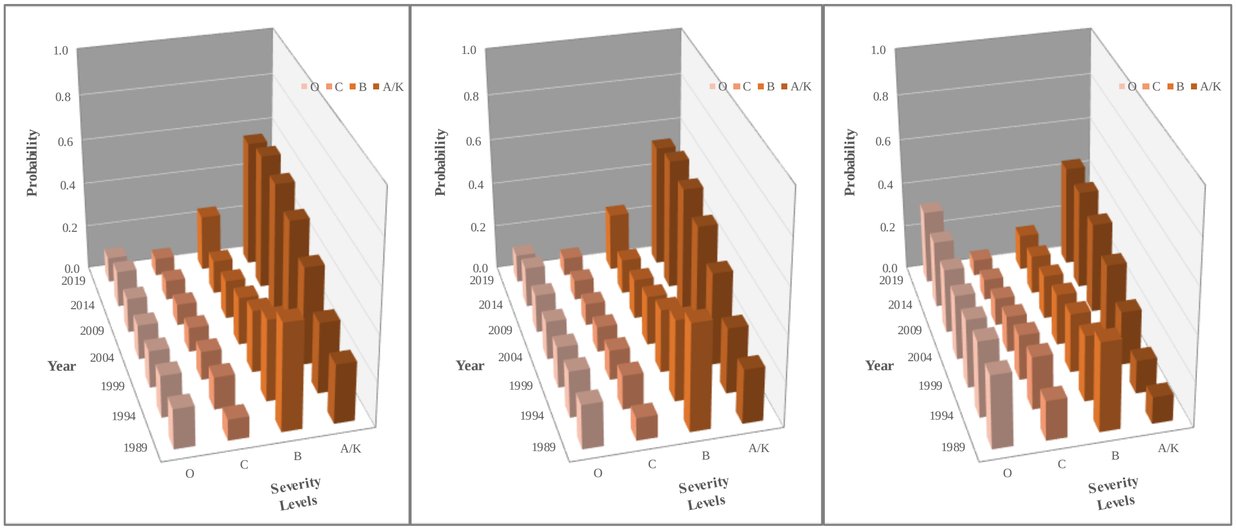
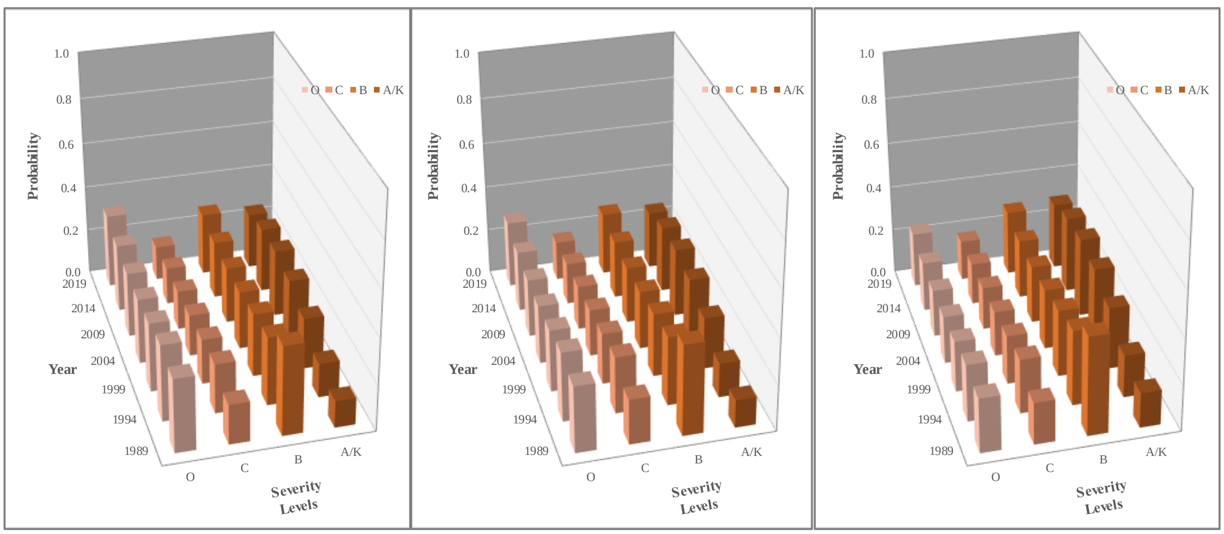


FIGURE 1 Probability Plots of Driver Injury Severity Levels across Different Years



(a) Driver condition 1 (b) Driver condition 2 (c) Driver condition 3



(d) Driver condition 4 (e) Driver condition 5 (f) Driver condition 6

O = No injury, C = Possible injury, B = non-incapacitating injury and A/K = incapacitating/fatal injury

FIGURE 2 Probability Plots of Driver Injury Severity Profiles across Different Years for Hypothetical Scenarios

TABLE 1 Sample Characteristics of Injury Severity Levels

Year	Injury severity levels					Total
	No injury	Possible injury	Non-incapacitating injury	Incapacitating injury	Fatal injury	
	Percentage (within year)					
1989	73.60	12.25	10.10	3.75	0.30	100.00
1994	71.30	15.15	8.35	4.95	0.25	100.00
1999	71.40	14.35	9.70	4.30	0.25	100.00
2004	68.05	12.75	10.55	7.90	0.75	100.00
2009	67.90	12.35	10.85	8.00	0.90	100.00
2014	71.70	13.05	10.15	4.65	0.45	100.00

TABLE 2 Crash Database Sample Statistics

Explanatory variables	Sample Share	
	Frequency	Percentage
Driver characteristics		
Driver gender		
Male	6,699	55.80
Female	5,300	44.20
Driver age		
Age less than 25	3,528	29.40
Age 25 to 64	7,452	62.10
Age more than 64	1,020	8.50
Influence of alcohol		
Under the influence of alcohol	696	5.80
No alcohol involved	11,304	94.20
Restraint system use		
Unrestrained	2,589	21.60
Restrained	9,411	78.40
Vehicle characteristics		
Vehicle type		
Pickup	1,596	13.30
Vans	756	6.30
Light truck	1,392	11.60
Utility	384	3.20
Passenger car	7,872	65.60
Roadway attributes		
Roadway class		
On interstate highway	935	7.80
Not on interstate highway	11,065	92.20
Road location		
Not intersection	4,008	33.40
Intersection	4,248	35.40
Intersection related	1,896	15.80
Driveway access	1,116	9.30
Other crash location	732	6.10
Traffic control device		
No traffic control	7,176	59.80
Traffic signal	3,096	25.80
Stop sign	948	7.90
Other traffic control device/sign	780	6.50
Environmental factors		

Day of week		
Weekdays	7,208	60.1
Weekends	4,792	39.90
Lighting conditions		
Day light	8,868	73.90
Dark-unlighted	948	7.90
Dark-lighted	1,776	14.80
Dawn-dusk	408	3.40
Road condition		
Dry	9,192	76.60
Wet	2,184	18.20
Other road condition	624	5.20
Crash characteristics		
Collision object		
Collision with moving vehicle	10,356	86.30
Collision with large object	1,356	11.30
Collision with other object	288	2.40
Manner of collision		
Rear-end	3,696	30.80
Head on	468	3.90
Angle	5,328	44.40
Sideswipe same direction	684	5.70
Other manner of collision	1,824	15.20

TABLE 3 Scaled Generalized Ordered Logit (SGOL) Model Results

Variables	Latent Propensity	Threshold between Possible and Non-Incapacitating Injury	Threshold between Non-incapacitating and Incapacitating/Fatal Injury
	Estimates (S.E.*)	Estimates (S.E.)	Estimates (S.E.)
Constant	1.452 (0.078)	-0.012 (0.047)	0.342 (0.066)
Temporal variables			
Time elapsed (linear)	0.051 (0.009)	-0.016 (0.006)	-0.055 (0.013)
Time elapsed (square)	-0.001 (0.0003)	---×	0.002 (0.0005)
Driver characteristics			
Driver gender (Base: Male)			
Female	0.324 (0.045)	---	---
Driver age (Base: Age 25 to 64)			
Age less than 25	-0.118 (0.061)	---	---
Age less than 25*Time elapsed (square)	-0.0003 (0.0002)	---	---
Age more than 64	0.164 (0.064)	-0.297 (0.090)	---
Under the influence of alcohol	0.673 (0.087)	-0.290 (0.103)	---
Restraint system use (Base: Restrained)			
Unrestrained*Time elapsed (linear)	---	0.026 (0.014)	---
Unrestrained*Time elapsed (square)	---	-0.0024 (0.001)	-0.0007 (0.0003)
Vehicle characteristics			
Vehicle type (Base: Car)			
Pickup	-0.415 (0.066)	-0.145 (0.085)	---
Vans	-0.168 (0.078)	0.254 (0.092)	---
Light Truck	-0.228 (0.124)	---	---
Utility	-0.267 (0.060)	---	---
Vehicle age	0.015 (0.003)	---	---
Roadway attributes			
Roadway class (base: Not on interstate highway)			
On interstate highway	0.405 (0.111)	---	---
On interstate highway*time elapsed (square)	-0.0006 (0.0003)	---	---
Road location (base: Not intersection and intersection)			
Intersection related	-0.217 (0.061)	0.137 (0.071)	0.212 (0.092)
Driveway access	-0.375 (0.071)	---	---
Other crash location	-0.241 (0.084)	---	---
Traffic control device (Base: Not stop sign)			
Stop sign	-0.133 (0.068)	---	---
Environmental factors			

Time of day (Base: Other than evening peak)			
Evening peak	-0.098 (0.044)	---	0.163 (0.067)
Day of week (Base: Weekday)			
Weekends	-0.069 (0.037)	---	---
Lighting condition (Base: Day Light)			
Dark-unlighted	0.145 (0.069)	-0.187 (0.097)	---
Road surface condition (Base: Dry)			
Wet	-0.094 (0.048)	---	---
Other road condition	-0.591 (0.103)	---	---
Crash characteristics			
Collision object (Based: Collision with moving vehicle)			
Collision with Large Object	0.739 (0.082)	-0.407 (0.081)	---
Manner of Collision (Base: Rear-end)			
Sideswipe same direction	-0.639 (0.113)	0.324 (0.127)	---
Head on	1.436 (0.178)	---	-0.294 (0.116)
Head-on*Time elapsed (linear)	-0.031 (0.010)	---	---
Angle	0.367 (0.049)	---	---
Scale parameter			
Time elapsed (linear)	0.0105 (2.088)	---	---

[*] S.E. = Standard Error

[×] Variable is not significant at 90% confidence level

TABLE 4 Elasticity Effects

Variables	Injury severity levels	Year						Mean Elasticity Effect	Confidence Interval	
		1989	1994	1999	2004	2009	2014		Lower Limit	Upper Limit
Female	No injury	-8.05	-9.46	-10.05	-11.16	-11.81	-11.29	-10.27	-10.99	-9.55
	Possible injury	19.17	18.22	18.67	18.27	18.73	21.50	19.05	19.01	19.09
	Non-incapacitating injury	25.61	25.31	25.81	25.72	26.18	29.47	26.41	26.27	26.54
	Incapacitating/Fatal injury	31.04	31.60	32.92	33.60	34.70	37.69	33.80	33.36	34.25
Under the influence of alcohol	No injury	-18.56	-21.56	-22.75	-25.21	-26.74	-26.05	-23.41	-24.99	-21.83
	Possible injury	10.29	5.79	5.88	2.97	2.50	7.64	5.77	4.55	7.00
	Non-incapacitating injury	88.07	83.53	81.48	78.05	77.20	89.12	82.85	81.39	84.30
	Incapacitating/Fatal injury	117.37	118.22	120.25	120.85	123.07	138.53	123.27	121.90	124.64
Unrestrained	No injury	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Possible injury	0.00	5.10	1.13	-11.85	-31.72	-53.04	-14.45	-19.88	-9.01
	Non-incapacitating injury	0.00	-5.46	-5.57	-2.57	-1.01	1.15	-2.14	-2.97	-1.31
	Incapacitating/Fatal injury	0.00	-3.36	5.43	26.09	57.61	113.59	35.89	29.92	41.85
Pickup	No injury	9.19	11.26	12.00	13.38	14.13	13.26	12.16	12.08	12.24
	Possible injury	-30.79	-31.59	-32.58	-32.46	-33.03	-35.53	-32.63	-32.69	-32.58
	Non-incapacitating injury	-21.11	-21.98	-22.97	-23.83	-25.02	-27.94	-23.93	-24.16	-23.71
	Incapacitating/Fatal injury	-24.75	-26.66	-28.31	-30.04	-31.84	-34.63	-29.81	-30.59	-29.02
On interstate highway	No injury	-10.69	-11.98	-11.08	-9.56	-6.11	-1.02	-8.43	-9.15	-7.71
	Possible injury	23.79	21.20	19.16	14.59	9.21	1.93	15.14	12.26	18.02
	Non-incapacitating injury	34.93	32.88	28.83	22.24	13.65	2.68	21.85	19.17	24.52
	Incapacitating/Fatal injury	44.83	43.41	38.83	30.53	18.64	3.46	28.57	26.18	30.96
Dark-unlighted	No injury	-3.66	-4.31	-4.56	-5.08	-5.40	-5.17	-4.68	-5.02	-4.35
	Possible injury	-5.83	-6.70	-6.63	-7.24	-7.33	-6.02	-6.64	-6.85	-6.43
	Non-incapacitating injury	25.28	24.29	23.75	22.59	22.08	24.30	23.67	23.27	24.08
	Incapacitating/Fatal injury	30.52	30.31	30.38	29.68	29.32	31.36	30.17	30.12	30.23
Collision with large object	No injury	-20.22	-23.60	-24.99	-27.58	-29.23	-28.56	-25.62	-25.76	-25.48
	Possible injury	3.24	-2.08	-2.02	-4.53	-4.40	-0.09	-1.72	-1.86	-1.59
	Non-incapacitating injury	102.95	97.42	95.50	90.92	90.14	103.60	96.67	96.51	96.83
	Incapacitating/Fatal injury	138.14	140.58	141.86	139.26	138.73	159.15	142.84	142.65	143.04
Head on	No injury	-42.62	-42.56	-39.05	-36.63	-32.16	-25.00	-36.41	-36.53	-36.28
	Possible injury	69.57	53.10	48.65	41.46	37.11	38.27	47.99	47.93	48.05
	Non-incapacitating injury	105.31	78.42	60.57	47.84	37.64	36.45	59.73	57.37	62.09
	Incapacitating/Fatal injury	358.73	282.17	235.45	196.57	167.85	158.80	220.34	197.14	243.55