INTRODUCING LATENT PSYCHOLOGICAL CONSTRUCTS IN INJURY SEVERITY MODELING: A MULTI-VEHICLE AND MULTI-OCCUPANT APPROACH

Patrícia Lavieri
The University of Texas at Austin
Department of Civil, Architectural and Environmental Engineering
301 E. Dean Keeton St. Stop C1761, Austin TX 78712
Phone: 512-471-4535, Fax: 512-475-8744, Email: laviepa@gmail.com

Chandra R. Bhat (*corresponding author*)
The University of Texas at Austin
Department of Civil, Architectural and Environmental Engineering
301 E. Dean Keeton St. Stop C1761, Austin TX 78712
Phone: 512-471-4535, Fax: 512-475-8744, Email: bhat@mail.utexas.edu

Ram M. Pendyala
Georgia Institute of Technology
School of Civil and Environmental Engineering
Mason Building, 790 Atlantic Drive, Atlanta, GA 30332-0355
Tel: 404-385-3754; Fax: 404-894-2278, Email: ram.pendyala@ce.gatech.edu

Venu M. Garikapati
Georgia Institute of Technology
School of Civil and Environmental Engineering
Mason Building, 790 Atlantic Drive, Atlanta, GA 30332-0355
Tel: 480-522-8067; Fax: 404-894-2278, Email: vgarikapati3@gatech.edu

Submitted for Presentation and Publication

Word count: 7,265 text + 4 tables/figures x 250 = 8,265 words
95th Annual Meeting of the Transportation Research Board

Committee on Safety Data, Analysis, and Evaluation (ANB20)

Revised November 15, 2015

ABSTRACT
This paper presents a comprehensive model of injury severity that accounts for unmeasured driver behavior attributes. Using indicators of risky and distracted/careless driving present in crash databases, the model system incorporates a latent variable component where latent constructs describing such behaviors can be modeled as a function of observed attributes of the driver. The model system also includes a measurement equation where the latent constructs of driver behavior are combined with other explanatory factors to model injury severity outcomes for all vehicular occupants in two-vehicle crashes. Building upon previous research, the paper presents a Generalized Heterogeneous Data Model (GHDM) capable of jointly modeling injury severity outcomes for all passengers in multiple vehicles by seat position. The model system is found to offer key insights on how various factors differentially affect injury outcomes for occupants in different seat positions. The results of the model have important implications for the design of safety interventions and advanced vehicular features and technologies. Engineering designs that accommodate the diminished capabilities of older drivers, include rear seat safety features, and alert drivers to frontal collisions before they occur (collision warning systems and automated braking systems) would contribute to substantial reductions in injury severity for various vehicular occupants.

Keywords: injury severity model, Generalized Heterogeneous Data Model (GHDM), latent variable modeling, driver behavior, safety interventions

1. INTRODUCTION

The World Health Organization (WHO) has reported that motor vehicle crashes are one of the most serious public health problems confronting both developed and developing countries around the world. Globally, the annual number of deaths on roadways is a staggering 1.24 million. The organization has stressed the need for a greater understanding of crash causation, injury severity and risky road-user behavior as important elements for preventing fatalities and injuries in the future (WHO, 2013). Although the number of crashes, injuries, and fatalities per million vehicle miles of travel is decreasing in the United States, the numbers are still large with more than 32,000 roadway fatalities in 2013. Another two million individuals were injured in roadway crashes (NHTSA, 2014).

Crashes involving passenger cars are of particular relevance because such crashes are associated with the majority of roadway deaths. More than one-half of the people that died in roadway crashes in 2013 were traveling in passenger cars (NHTSA, 2014). Among the crashes involving a fatality, two-thirds involve a single vehicle and one-third involves multiple vehicles. Among the non-fatal injury crashes, two-thirds of the crashes involve multiple vehicles and one-third involve a single vehicle. There were 305,000 two-vehicle injury crashes in the United States in 2013 (NHTSA, 2014).

Despite considerable research devoted to crash data analysis and injury severity modeling, there is a paucity of literature devoted to fully accounting for driver behavior in injury severity models. The aim of this paper is to fill a critical gap in the literature by presenting a model system that captures the impact of driver's behavior on the injury severities of crash victims. In addition, while prior research has generally focused on the injury severity of the most injured victim, the model system in the current study *jointly* models the injury severity of all vehicle occupants associating them to their respective seat positions in the vehicle(s) involved, and captures the cross-effects of the driver characteristics of one vehicle over to the other vehicle (in a multi-vehicle crash). By doing so, the model provides valuable insights on the vulnerability of passengers in various seat positions, thus helping to identify safety interventions and engineering designs that improve safety outcomes for all passengers regardless of seating position. The model system also accounts for the endogeneity of seat-belt use and alcohol consumption; inherently safe drivers are likely to use seat belts and avoid driving under the influence – accounting for such self-selection is critical to modeling the effects of various explanatory factors on injury severity.

Driver behavior is usually not adequately considered in injury severity models because of data limitations; crash injury severity data generally do not include behavioral characteristics and include virtually no psychological measurements. Data on driving behavior is often self-reported, and may not be available in the context of specific crashes and injury severity outcomes. Recent methodological enhancements, however, have made it possible to account for unobserved heterogeneity in the driver population, as well as endogeneity in driver behavior. Such methodological approaches capture the variation in driver behavior (due to unobserved psychological factors) that is prevalent on the roadway system. The Integrated Choice and Latent Variable (ICLV) modeling methodology and the Generalized Heterogeneous Data Model (GHDM) offer the econometric tools necessary to incorporate latent driver behavior constructs in injury severity models. Rather than using exogenous variables to characterize attitudes and psychological characteristics, endogenous variables in the crash data set (such as the use of seat belt and alcohol involvement prevalent in police crash reports) can be used to develop latent constructs of driver behavior.

In addition to incorporating unobserved heterogeneity in the driver population in injury severity models through the use of latent constructs, this study contributes to the literature by presenting a model system that jointly models the injury severity of all people involved in a crash. This has not been done previously due to methodological constraints; however, the methodologies invoked in this paper are capable of accounting for injury severity across all vehicle occupants. Finally, the proposed framework makes it possible to accommodate endogeneity of specific factors in injury severity models in an easy and flexible manner.

The remainder of this paper is organized as follows. The next section provides an overview of the literature on crash and injury modeling. The third section offers an overview of the modeling framework. The fourth section provides an overview of the modeling methodology, while the fifth section describes the data used in this research. The sixth section presents the model estimation results. Conclusions and implications of the findings are in the seventh and final section of the paper.

2. MODELING INJURY SEVERITY

Driver behavior is a critical determinant of road traffic crashes (Petridou and Moustaki, 2000), and it is therefore of much interest and importance to account for this dimension in the study and modeling of crash occurrence, crash type, and injury severity. There is a substantial body of literature that attempts to link individual personality traits to driving behavior and the likelihood of committing traffic violations and being involved in traffic accidents (recent examples include Constantinou et al, 2011; Taubman-Ben-Ari and Yehiel, 2012; Zhao et al, 2013; and Beanland et al, 2014). However, this stream of research has not explicitly linked driver behavior to the analysis injury severity. The reason for this disconnect is that studies of driver behavior and studies of injury severity outcomes use different streams of data. Most driving behavior studies rely on self-reported data as this is generally the most cost-effective manner of measuring personality traits (through attitudinal, perception, and self-assessment statements) and behavior (e.g., frequency of actions such as passing, tail-gating, speeding) simultaneously (Beanland et al, 2014). This data collection approach suffers from the limitation that it relies on the respondent’s memory and judgement, and therefore limits the amount of information that can be obtained about the relationship between situational contexts and driving behaviors, as well as the consequent outcomes. The approach therefore makes it difficult to explicitly connect driver personality and behavior with crash injury severity outcomes. Despite data limitations, the field of crash injury severity modeling has seen important methodological and empirical developments. Savolainen et al (2011) provide a review of statistical methods used to model crash injury and severity. More recently, Mannering and Bhat (2014) discussed methodological frontiers in accident research.

Studies of injury severity have generally treated the injury severity variable as either binary (injury versus non-injury) or as a multiple response variable (no apparent injury, possible injury, minor injury, serious injury, and fatal injury). These outcomes may be treated as ordered or unordered outcomes. Savolainen et al (2011) group studies according to their approach and find that ordered probit models are the most frequently used, followed by multinomial logit, nested logit, and mixed logit respectively. Yasmin and Eluru (2013) performed an empirical comparison of ordered response and unordered response models in the context of driver injury severity. They find that an ordered system that allows for exogenous variable effects to vary across alternatives and accommodates unobserved heterogeneity offers almost equivalent results to that of the corresponding unordered system.

A study that explicitly uses a psychological construct to moderate the impact of other variables on injury severity outcomes is that by Paleti et al (2010). They jointly model, through a two-equation system, the driving behavior (aggressive driving) and injury severity propensity. Donmez and Liu (2015) consider different types of distracted driving as explanatory variables in a simple single equation ordered logit model of injury severity. This study aims to contribute to the literature by exploiting recent econometric methodological advances to model injury severity outcomes while accounting for unobserved driver characteristics. Endogenous variables (available in police reports, such as seat belt use and alcohol involvement) may be used as indicators of latent variables; latent constructs developed using these variables can then be introduced in the injury severity model through a simultaneous equations model framework.

Another area where this study makes a contribution is the modeling of injury severity for all individuals involved in a crash. Most injury severity studies model only the injury of one individual in the crash – the vehicle's driver or the most severely injured occupant. Kim et al, (2013), and Donmez and Liu (2015) are examples of studies that focus on the injury severity of the driver, while Castro et al (2013) and Weiss et al (2014) model the injury severity of the most severely injured occupant. When only one person is modeled, important information that could be used to guide comprehensive traffic safety measures and technologies is missed. This occurs because although the injury severity of individuals involved in a crash are likely to be correlated (which necessitates the joint modeling framework in this paper), it will generally not be true that the injury severity of the most severely injured person in a crash is quite representative of the injury severity sustained by other individuals involved in a crash. Indeed, studies that used seat position as an explanatory variable or that modeled risk ratios between front and rear seats did identify significant differences on injury levels of passengers seating in the front seats compared to those seating in the rear-seats (e.g. Mayrose and Priya, 2008; Durbin et al., 2015). At the same time, modeling injury severities independently based on seat position, as done in many earlier studies just mentioned, would also be inefficient because of the correlation in injury severities across seat positions in the same crash; that is, unobserved factors specific to a crash and specific to a vehicle may simultaneously increase or decrease the injury severity sustained by each individual in each seat position relative to their peers in the corresponding seat positions in other crashes and other vehicles.

Studies that have examined injury severity of multiple occupants are prevalent in the literature. Yasmin et al (2014) modeled the injury severity of the two drivers involved in a two-vehicle crash. Abay et al (2013) also examined the injury severity of the two drivers involved in a crash using a multivariate probit model. One study that considers injury severity of all occupants is that by Eluru et al (2010) who employed a copula-based approach to model the multiple occupant injury phenomenon. They find correlated unobserved factors in injury outcomes across occupants and recommend that crash studies adopt approaches in which injury outcomes are modeled simultaneously across all vehicle occupants.

Some variables used to explain injury severity are treated as exogenous when in fact they should be treated as endogenous. Factors that influence these variables may also influence the severity of injuries sustained in a crash, rendering such variables correlated with the unexplained part (error term in the model) of injury severity. Examples of such variables include the decision to wear a seat belt, the decision to drive while under the influence or impaired, and the decision to acquire a vehicle with special safety features. Personality traits and intrinsic driver behavior characteristics that make an individual wear a seat belt, purchase a vehicle with special safety features, and avoid driving while impaired are also likely to impact injury severity outcomes as such drivers are likely to be inherently safer and more risk-averse in their operation of a vehicle. Eluru and Bhat (2007) and Abay et al (2013) are examples of studies that consider the endogeneity of seat belt use; not accounting for endogeneity may lead to an overestimation or underestimation of the effect of the corresponding variables on injury severity outcomes. For example, a driver in a vehicle with enhanced safety features may compensate for their presence by engaging in risky driving behaviors (speeding, for example); if the variables describing safety features are treated as exogenous variables, their potential (beneficial) impacts will be under-estimated.

This paper makes a methodological contribution by exploiting advanced econometric tools to account for three empirical considerations in injury severity modeling. First, the paper introduces latent constructs to capture unobserved driver behavior and psychological traits; second, the methodology jointly models the injury severity outcomes for all vehicle occupants involved in a crash; and third, the methodology accounts for endogeneity of safety related variables.

3. CONCEPTUAL FRAMEWORK

The conceptual framework adopted for the modeling effort in this paper is an adaptation of the idea of a contextual model proposed by Sümer (2003) in which a vehicular crash is viewed as a consequence of both distal context and proximal context variables. The proximal context mediates the impact of the distal context on the outcome. Figure 1 presents an overview of the conceptual framework. In the modeling framework of this paper, the injury severities are a consequence of both distal and proximal contexts of both drivers involved in the crash.

The *distal* context variables in the modeling framework include characteristics of the driver that are inherent to the individual such as age, gender, and personality traits. These characteristics are relatively stable across time and not specific to the crash circumstances. On the other hand, *proximal* context variables include both stable and transitory factors closely related to the crash. Proximal variables include driver behavior (e.g., wearing seat belt, speeding), environmental conditions, roadway characteristics and condition, vehicle characteristics (which may also be viewed as distal context variables depending on the circumstances of the crash), and crash outcome variables. As the injury severity of all vehicle occupants is being modeled, age and gender of passengers are also proximal variables because they are not directly related to the driver and can change from one trip to the next. Similarly, presence of children is also a proximal variable.

This study is limited to an analysis of two-vehicle crashes. As such, one vehicle is part of the proximal context of the other. Some variables describing the proximal context of one vehicle also impact the other vehicle, representing a reciprocal effect. The behavior of the driver in one vehicle can impact injury severity of occupants in the other vehicle. The driver behaviors represented in the model include *distracted/careless driving* behavior and *risky driving* behavior. They are assumed to be a consequence of distal factors, although some proximal factors such as presence of passengers (or the interaction of proximal and distal factors) may also affect driving behavior constructs.

Psychological traits are assumed to impact driving behaviors. However, such traits are not observed or measured in crash databases. The framework requires a minimum of two indicators for identification purposes, but can accommodate as many as desired. The proposed framework is quite flexible and may be used for crashes involving multiple vehicles (more than two vehicles) of any type. As shown in the framework, injury severities of all vehicle occupants in two-vehicle crashes are modeled and psychological constructs describing both drivers involved in the crash are considered to have influence on these severities. The occupants of the vehicles are linked to their seat positions, resulting in five possible injury severity outcomes associated with: 1) the driver's seat; 2) the front passenger seat; 3) the back left seat; 4) the back middle seat; and 5) the back right seat. Each driver has two latent constructs – risky driving behavior and careless/distracted driving behavior. Both the latent constructs from each driver affect the injury severities of all occupants in both vehicles. Each person in a vehicle is affected equally by the same construct corresponding to a specific driver. The modeling framework accommodates the effects of: 1) risky behavior of the driver on his/her own vehicle; 2) risky behavior of the driver on the other vehicle; 3) careless behavior of the driver on his/her own vehicle; and 4) careless behavior of the driver on the other vehicle.

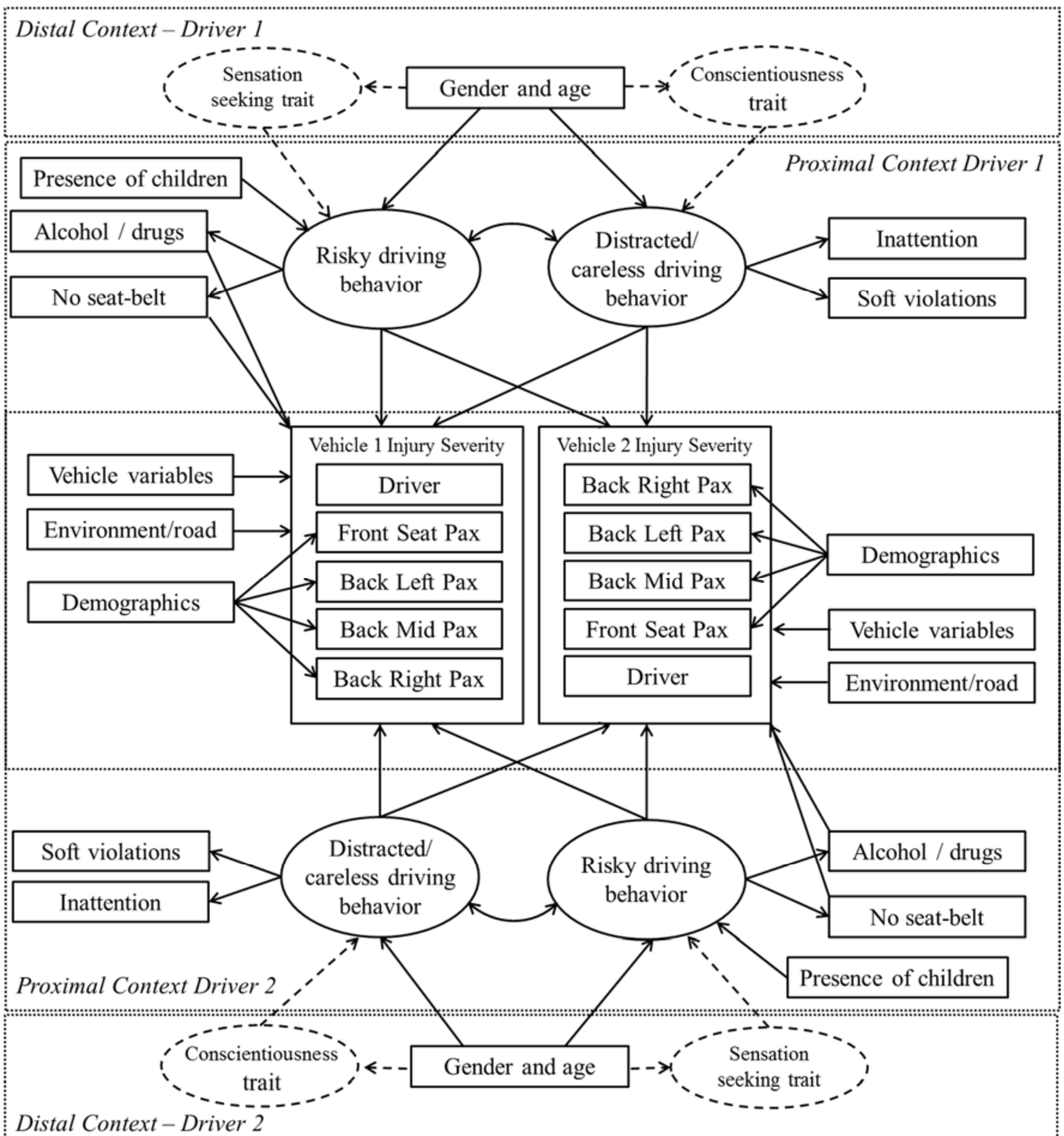

FIGURE 1 Conceptual framework of injury severity model system

For the empirical application conducted in this paper, the outcomes associated with careless/distracted behavior are soft violations and inattention, while the outcomes associated with risky behavior are no seat belt use and alcohol impairment (the endogeneity of these last two variables is also represented).

An issue that arises in this modeling effort is that the labeling of drivers, as driver 1 and driver 2, is arbitrary and does not represent any real distinction between types of drivers. To address this labeling issue, the loadings of the latent factors on each binary outcome are constrained to be the same across drivers. Additionally, because the effect of the demographics on the latent variables (careless/distracted driving and risky driving) should also be invariant to the labeling of a driver as 1 or 2, the loading of driver demographics on latent characteristics are held to be constant across the two drivers (that is, a single relationship holds between driver demographics and driver latent constructs). The same follows for the correlation between risky and distracted/careless driving behavior latent variables, which should be a unique parameter. The parameters associated with exogenous explanatory variables (environment, road condition, crash type) on injury severity are also constrained to be the same across the occupants of the two vehicles seated in the same position, because the vehicles are also labeled arbitrarily. For the same reason, thresholds associated with the propensity of individuals seated in the same position in a vehicle to experience injury severity of different levels are constrained to be the same across vehicles. Both injury severities and the binary outcomes are coded as ordinal variables. The injury severity levels are: 1) no apparent injury; 2) possible injury; 3) minor injury; and 4) serious or fatal injury. The binary outcomes take on a value of one or two for notational consistency.

4. MODELING METHODOLOGY

A special case of the GHDM approach proposed by Bhat (2015) is used for the modeling effort in this paper. It constitutes a special case because all of the outcomes in this study are ordinal, thus avoiding the necessity to deal with a mixture of dependent variable types. The model system is composed of a *structural equation* component and a *measurement equation* component. In the structural equation, driver age and gender, and presence of children in the vehicle are used to explain distracted/careless driving behavior and risky driving behavior. Correlation across these two latent constructs is accommodated in the model formulation. In the measurement equation, the two latent variables are loaded on the four indicators (two for distracted/careless driving, and two for risky driving) and also on the injury severity of every vehicle occupant. All of the other explanatory variables are also loaded on the injury severity outcomes, and interactions between the explanatory variables and the latent variables can be accommodated in the model specification.

4.1 Latent Variable Structural Equation Model Component

Consider the latent variable z_l^* and write it as a linear function of covariates:

$$z_{qvl}^* = \boldsymbol{\alpha}_l' \boldsymbol{w}_{qv} + \eta_{qvl} \tag{1}$$

$\boldsymbol{w}_{qv}$ is a vector of observed covariates (excluding a constant) corresponding to crash $\boldsymbol{q}$ and vehicle driver $\boldsymbol{v}$, $\boldsymbol{\alpha}_l$ is a corresponding $(D \times 1)$ vector of coefficients (note that we mantain the same coefficient vector $\boldsymbol{\alpha}_l$ across drivers $\boldsymbol{q}$ because of the arbitrary labeling issue discussed earlier), and η_{qvl} is a random error term assumed to be standard normally distributed for identification purposes (see Stapleton, 1978). Next, define the $(L \times D)$ matrix $\boldsymbol{\alpha} = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, ..., \boldsymbol{\alpha}_L)'$, the $(L \times 1)$ vectors

$\boldsymbol{z}^*_{qv} = (z^*_{qv1}, z^*_{qv2}, ..., z^*_{qvL})'$ and the $(L \times 1)$ vector $\boldsymbol{\eta}_{qv} = (\eta_{qv1}, \eta_{qv2}, \eta_{qv3}, \ldots, \eta_{qvL})'$. The $\boldsymbol{\eta}_{qv}$ vector is distributed L-variate standard normal as follows: $\boldsymbol{\eta}_{qv} \sim MVN_L[\boldsymbol{0}_L, \boldsymbol{\Gamma}]$, where $\boldsymbol{0}_L$ is an $(L \times 1)$ column vector of zeros, and $\boldsymbol{\Gamma}$ is a $(L \times L)$ correlation matrix. In the empirical analysis in this paper $L = 2$ (two latent variables) and $\boldsymbol{\Gamma} = \begin{bmatrix} \mathbf{1} & \sigma_{\mathbf{12}} \\ \sigma_{\mathbf{12}} & \mathbf{1} \end{bmatrix}$, where $\sigma_{\mathbf{12}}$ represents the correlation between the latent variables for each vehicle driver $\boldsymbol{v}$ in crash $\boldsymbol{q}$ (ie.: in Figure 1, the correlations between distracted driving and risky driving behavior of driver $\boldsymbol{v}$ in crash $\boldsymbol{q}$; we expect $\sigma_{\mathbf{12}}$ to be positive because drivers who are generally more distracted relative to their observationally equivalent peers should be more likely to exhibit more risky driving behavior than their observationally equivalent peers). The reader will note that this driver-specific correlation is invariant across drivers in a crash and drivers across crashes.

In matrix form, we may write Equation 1 as

$$z^*_{qv} = \boldsymbol{\alpha}' \boldsymbol{w}_{qv} + \eta_{qv} \tag{2}$$

with the parameters to be estimated in the structural equation being $\boldsymbol{\alpha}$ and the non-diagonal element of $\boldsymbol{\Gamma}$.

4.2 Latent Variable Measurement Equation Model Component

Let $\boldsymbol{n}$ be the index of the ordinal outcomes associated with each vehicle and each crash. The number of outcomes will technically vary across vehicle-crash conditions (because the number of occupants will vary across vehicles and crashes, and injury severities of vehicle occupants constitute a subset of ordinal outcomes corresponding to each vehicle and each crash). But, for presentation and programming ease, we will maintain a fixed number of ordinal outcomes for each vehicle-crash combination and place null vectors in the $\boldsymbol{x}_{qvn}$ vector (corresponding to the covariate vector specific to ordinal outcome $\boldsymbol{n}$ of vehicle $\boldsymbol{v}$ and crash $\boldsymbol{q}$) for those vehicle-crash combinations for which a specific outcome $\boldsymbol{n}$ is not relevant (for example, if there is no occupant in the front passenger seat of a vehicle involved in a crash, the $\boldsymbol{x}_{qvn}$ vector is the null vector for $\boldsymbol{n}$ corresponding to the injury severity of the person in this seat). The measurement equation may be written as

$$y^*_{qvn} = \boldsymbol{\gamma}'_n \boldsymbol{x}_{qvn} + \boldsymbol{d}'_n \boldsymbol{z}_{qv}{}^* + \varepsilon_{qvn} \tag{3}$$

where ε_{qvn} is the standard normal random error vector for the n^{th} ordinal outcome which is assumed to be independent across outcomes $\boldsymbol{n}$ (though there is covariance across the y^*_{qvn} variables for the $\boldsymbol{n}$ outcomes because of the presence of the $\boldsymbol{z}_{qv}{}^*$ vector). What we observe for each outcome is the ordinal category of the outcome (for example, in the context of seat belt use, there are only two categories –yes or no –while, in the context of injury severity there are four categories : no apparent injury, possible injury, minor injury, serious or fatal injury). If the observed outcome for the n^{th} ordinal outcome is $\boldsymbol{a}_n$, then in the ordered-response formulation, this implies that $\boldsymbol{\psi}_{n,a_n-1} < \boldsymbol{y}_{qvn}{}^* < \boldsymbol{\psi}_{n,a_n}$ where $\psi_{n,0} < \psi_{n,1} < \psi_{n,2} ... < \psi_{n,J_n-1} < \psi_{n,J_n}$; $\psi_{n,0} = -\infty$, $\psi_{n,1} = 0$, and $\psi_{n,J_n} = +\infty$ (note that for the binary outcomes $J_n = 2$, and there are no threshold to be estimated. Note that J_n represents the number of categories of the ordinal outcome $\boldsymbol{n}$). The parameters to be estimated in the measurement equation are for each outcome $\boldsymbol{n}$, the $\boldsymbol{\gamma}_n$ parameters on observed covariates, the $\boldsymbol{d}_n$ parameters representing the loadings of the latent variables for each vahicle driver-crash combination on the outcomes corresponding to that vehicle-crash combination, and the ψ thresholds.

Readers are referred to Bhat (2015) for a detailed discussion on identification issues and the detailed estimation approach. The model system uses the features of the GHDM to accommodate correlation across both vehicles and all occupants involved in a crash. However, different from previous applications of the GHDM and previous injury severity studies, the model proposed in this paper offers a versatile structure that accommodates cross-effects between vehicles through mapping matrices. The mapping matrices are both used to solve the arbitrary labeling issues noted previously and to accommodate cross-vehicle effects. The mapping matrices can be easily expanded to accommodate additional vehicles, additional occupants or seating positions, and additional latent and endogenous variables.

5. DATA DESCRIPTION

The data used in this study is derived from the latest wave (2013) of the National Automotive Sampling System (NASS) General Estimates System (GES) crash database. The GES crash database provides data on a representative sample of crashes of all types involving all types of vehicles. The analysis and modeling effort is limited to crashes involving two passenger vehicles. Table 1 presents an overview of the descriptive characteristics of the dataset. The cleaned data set used for model estimation included 3,429 crashes. These crashes involve 9,177 individuals – 6,858 drivers and 2,319 passengers. The vehicles have up to four occupants (the few observations with more than four occupants in a vehicle had missing values and were removed). The variable indicating whether an airbag was deployed or not could not be included in the model specification because of the large prevalence of missing values for this variable. The crashes included in the estimation data set were limited to those involving "automobiles" as defined in the GES analytical user's manual. Due to a high prevalence of missing values for several driver behavior indicators (e.g., if driver was speeding, different types of violations, reckless driving, use of cell phone, distractions inside or outside vehicle), the set of indicators was limited to the following where complete data was consistently available:

1) For risky driving behavior
    a. Alcohol or drug use
    b. Non seat belt use
2) For distracted/careless driving behavior
    a. Inattention
    b. Soft violations that can be associated with a distraction (fail to yield, fail to stop, improper turn, improper use of lane, fail to obey sign or signal)

A large percent of crashes occur in the midday (9AM to 4PM) in the daylight hours, simply because there is more travel during those periods. Similarly, most accidents occur in clear weather (72.2 percent). Very few crashes are associated with roadways with very high speed limits of 70-85 mph presumably because there are fewer roadways (and hence less travel) with such speed limits. Nearly 60 percent of crashes occur at intersections where there are multiple conflict points. With respect to driving behaviors, soft violations are involved in 16 percent of the crashes. Risky behaviors are involved in small percent of crashes (about five percent or less). There is no apparent injury in two-thirds of the crashes. Owing to the high prevalence of missing values for crashes on important endogenous outcomes, as well as the aggregate nature of the weights in GES, it was considered prudent to use the unweighted sample for model estimation.

TABLE 1 Descriptive characteristics of the crash database sample

Person Variables		
Drivers (6858 observations)		
Female	3669	53.50%
Male	3189	46.50%
Age 16 to 24	1916	27.94%
Age 25 to 35	1588	23.16%
Age 36 to 45	1421	20.72%
Age 46 to 65	1178	17.18%
Age > 65	755	11.01%
Alcohol/drugs use	165	2.41%
No Seat-belt use	127	1.85%
Inattention	370	5.40%
Soft violations	1117	16.29%
Passengers (2319 observations)		
Female	1329	57.31%
Male	990	42.69%
Age < 15	706	30.44%
Age 15 to 24	642	27.68%
Age 25 to 35	354	15.27%
Age 36 to 65	429	18.50%
Age > 65	188	8.11%
Vehicle Variables		
Vehicle type (6858 observations)		
Sedan	5151	75.11%
Hatchback	393	5.73%
Station Wagon	537	7.83%
Convertible	128	1.87%
Others	649	9.46%
Vehicle age in years (6858 observations)		
≤ 5	2315	33.76%
6 to10	2151	31.36%
> 10	2392	34.88%
Area of impact (6858 observations)		
Front	5074	73.99%
Left	400	5.83%
Right	482	7.03%
Back	902	13.15%

Crash Variables		
Collision type (3429 observations)		
Rear-end	1269	37.01%
Frontal	270	7.87%
Angle	1499	43.72%
Side: same direction	307	8.95%
Side: opposite direction	62	1.81%
Other	22	0.64%
Speed limit (3429 observations)		
≤ 35 mph	1642	47.89%
> 35 mph	1787	52.11%
Junction type (3429 observations)		
Intersection	2047	59.70%
Access	424	12.37%
Other type of junction	874	25.49%
Not a junction	84	2.45%
Time of the day (3429 observations)		
12am to 6am	208	6.07%
6am to 12am	3221	93.93%
Light conditions (3429 observations)		
Daylight	2544	74.19%
Dawn or dusk	125	3.65%
Dark	195	5.69%
Dark with artificial light	565	16.48%
Weather conditions (3429 observations)		
Clear	2474	72.15%
Rain	335	9.77%
Snowing	52	1.52%
Other	568	16.56%
Injury Severity		
Vehicle occupants (9177 observations)		
No apparent injury	6107	66.55%
Possible injury	1281	13.96%
Minor injury	1148	12.51%
Serious/fatal injury	641	6.98%

6. MODEL ESTIMATION RESULTS

Model estimation was undertaken for all occupants jointly, accounting for correlation among unobserved factors through the two latent variables. The model structure also accommodated cross-effects where the behavior each driver affects outcomes for both vehicles involved in the crash. A variety of model specifications were tested treating explanatory variables as both alternative specific and generic in nature; for some variables, such as light conditions, it would not be reasonable to test for different coefficients across the seat positions and hence such variables were treated as generic variables. Other variables, such as side of impact, were tested to determine whether a generic treatment would be appropriate. In general, the limitations of the data set, including missing data on a number of key indicators of distracted driving (e.g., cell phone use, texting) prevented the full exploitation of the capabilities of the model formulation.

6.1 Results of the Structural Equation Component

The top half of Table 2 presents results of the structural equation component of the model system. With respect to distracted and careless behavior, the results indicate that females are less likely to be distracted and careless. The literature (e.g., McEvoy et al, 2007) suggests that there may be reasons for both males and females to be more distracted than the other. Males tend to be more distracted by outside distractions and mobile phone use than female drivers, while females are more likely to talk to other passengers while driving.

TABLE 2 Results of the structural equation component and four binary outcomes of the measurement equation component

Structural Equation Model		
Variables	**Coefficient**	**t-stat**
Driver's risky behavior		
Female	-0.7922	-21.05
Presence of children in the vehicle	-0.3578	-9.38
Age 26-35 (base15-25 years old)	-0.3344	-11.65
Age 36-65 years old	-0.5124	-14.85
Age > 65 years old	-0.6439	-14.15
Driver's distracted/careless behavior		
Female	-0.0815	-6.04
Age > 65 (base is less or equal to 65 years old)	0.0502	2.87
Correlation between risky and distracted/careless behaviors	0.2600	2.10
Measurement Equation - Latent Variable Loadings on the Binary Outcomes		
No Seat Belt Use		
Constant – no seat belt use	-2.0057	-5.39
Risky driving behavior	0.3866	4.92
Alcohol Use		
Constant – alcohol use	-1.9613	-2.46
Risky driving behavior	0.6055	7.47
Inattention		
Constant – inattention	-1.6039	-61.06
Distracted/careless driving behavior	0.0655	9.34
Soft Violations		
Constant – soft violations	-0.9891	-33.30
Distracted/careless driving behavior	0.1776	3.91

Those older than 65 years of age are more likely to be distracted and careless. One possible reason for this is that aging is related to an increase in both visual impairment and difficulty in dividing attention between driving and any other activity (Owsley et al, 1998). Being female, being older and the presence of children in the vehicle are all negatively associated with risky driving behavior. These results are consistent with those reported in the literature (for example, Paleti et al, 2010) suggesting that male and younger drivers are more likely to partake in aggressive driving acts than female and older drivers respectively, while Fleiter et al (2010) notes that drivers are more careful when children are present. Finally, the correlation between risky driving behavior and distracted/careless driving behavior is, as expected, positive and statistically significant.

6.2 Results of the Measurement Equation Component

The bottom half of Table 2 presents results for the binary outcome variables in the measurement equation component of the model system. The four binary outcome variables include no seat belt use, alcohol use, inattention, and soft violations. Each binary outcome variable equation includes a constant and a latent variable (risky driving behavior or distracted/careless driving behavior) on the right hand side. The negative constants suggest that drivers generally tend to be safe and alert. As expected, risky driving behavior is positively associated with no seat belt use and alcohol involvement. Likewise, distracted and careless driving behavior is positively associated with inattention and commission of soft violations.

The measurement equation component also includes an extensive set of explanatory variables and latent factors to capture the influence of various attributes on the injury severity of occupants seated in different positions. The model estimation results for the injury severity component of the measurement equation are presented in Table 3. In addition to the latent constructs, the model includes a number of occupant characteristics (age and gender), vehicle characteristics (vehicle type and age), crash characteristics (collision type, area of impact), environmental variables (time of day, light conditions, and weather conditions), and roadway characteristic variables (speed limit, intersection type, trafficway descriptors).

Males have a lower propensity to sustain severe injuries when compared to females in all seat positions, consistent with findings reported by Eluru et al (2010). Children 14 years of age or younger are less prone to severe injuries in all back seat positions, reinforcing the adage that children are safest when in the rear seat. Those older than 65 years of age are more susceptible to severe injuries in all seat positions, an indication the weakened physical state at an advanced age. The absence of seat belt use contributes significantly to severe injury outcomes, reaffirming that seat belts can reduce the impact of crashes on vehicle occupants.

Occupants are more likely to sustain severe injuries when seated in hatchbacks and convertibles (as opposed to sedans and station wagons, that are likely larger and safer vehicles), a finding consistent with that reported by Ju and Sohn (2011). Compared to newer vehicles, occupants are likely to sustain injuries in older vehicles with the highest propensity for severe injuries in vehicles over 10 years of age. The condition of the vehicles and the likelihood that vehicles of such vintage do not include the latest safety features contribute to this finding (Bilston et al, 2010). Both the absence of seat belt use and alcohol impairment contribute significantly to severe injury outcomes even after accounting for their endogeneity, a finding that is consistent with expectations.

Rear-end crashes are associated with less severe injuries while frontal collisions result in more severe injuries across all seating positions. Older individuals greater than 65 years of age are likely to sustain more serious injuries when in a side-impact crash (compared to younger counterparts). In terms of the environmental conditions, crashes occurring in the overnight hours of 12AM to 6AM are most likely to result in severe injuries, possibly due to excessive speeding (when there is no traffic on the roadways), darkness, and impaired driving. Both darkness and dawn/dusk hours are associated with more severe injury outcomes compared to daylight conditions or dark-with artificial light conditions, consistent with expectation.

TABLE 3 Injury severity propensity estimates

Variable name	Driver		Front Passenger		Back left seat		Back middle seat		Back right seat	
	Coef	t-stat	Coef	t-stat	Coef	t-stat	Coef	t-stat	Coef	t-stat
Constant	-0.8675	-8.15	-0.4064	-3.06	-0.2413	-3.31	-0.1417	-4.11	-0.2343	-3.4
Threshold parameters										
Threshold 1	0.7802	30.12	0.9165	17.33	1.1803	7.47	1.0515	3.91	0.9199	8.00
Threshold 2	1.9356	44.50	2.0559	25.65	2.5991	9.34	2.3579	4.92	2.6204	11.02
Occupant Characteristics										
Male	-	-	-0.424	-4.39	-0.424	-4.39	-0.424	-4.39	-0.424	-4.39
Age (base: 15-65 years old)										
0-14	-	-	-	-	-0.5318	-3.69	-0.5318	-3.69	-0.5318	-3.69
>65	-	-	0.6383	3.41	0.6383	3.41	0.6383	3.41	0.6383	3.41
No seat-belt use (base: seat-belt use)	1.7548	3.07	1.7548	3.07	1.1847	3.07	1.1847	3.07	1.1847	3.07
Driver alcohol use	0.6351	3.09	0.6351	3.09	0.6351	3.09	0.6351	3.09	0.6351	3.09
Vehicle Characteristics										
Vehicle type (base: sedan and station wagon)										
Hatchback or convertible	0.2325	3.36	0.2325	3.36	0.2325	3.36	0.2325	3.36	0.2325	3.36
Vehicle age (base: < 5 years)										
Vehicle age between 5 and 10 years	0.0704	3.83	0.0704	3.83	0.5159	2.21	0.5159	2.21	0.5159	2.21
Vehicle age more than 10 years	0.2389	3.98	0.2389	3.98	0.5861	2.56	0.5861	2.56	0.5861	2.56
Road Variables										
Speed limit (base is > 35 mph)										
< 35 mph	-0.3628	-6.00	-0.3628	-6.00	-0.3628	-6.00	-0.3628	-6.00	-0.3628	-6.00
Junction type (base: intersection)										
Access or not a junction	0.1688	2.66	0.1688	2.66	0.1688	2.66	0.1688	2.66	0.1688	2.66
Other type of junction	0.6130	3.05	0.6130	3.05	0.6130	3.05	0.6130	3.05	0.6130	3.05

TABLE 3 Injury severity propensity estimates (continued)

Variable name	Driver		Front Passenger		Back left seat		Back middle seat		Back right seat	
	Coef	t-stat	Coef	t-stat	Coef	t-stat	Coef	t-stat	Coef	t-stat
Crash Characteristics										
Collision type (base: angle collision)										
Rear-end	-0.7051	-10.27	-0.7051	-10.27	-0.7051	-10.27	-0.7051	-10.27	-0.7051	-10.27
Frontal	0.9509	8.87	0.9509	8.87	0.9509	8.87	0.9509	8.87	0.9509	8.87
Side: same direction	-1.4548	-10.90	-1.4548	-10.9	-1.4548	-10.9	-1.4548	-10.9	-1.4548	-10.9
Side: opposite direction	-0.5010	-2.23	-0.5010	-2.23	-0.5010	-2.23	-0.5010	-2.23	-0.5010	-2.23
Area of impact on each vehicle (base: front)										
Left	0.5176	4.35	-	-	0.5176	4.35	-	-	-	-
Right	-	-	0.576	3.06	-	-	-	-	0.576	3.06
Back	-0.2592	-2.93	-0.2592	-2.93	-0.2592	-2.93	-0.2592	-2.93	-0.2592	-2.93
Side impact × elder passenger	-	-	0.1325	4.42	0.1325	4.42	0.1325	4.42	0.1325	4.42
Environment										
Time of the day (base: 6am to 12am)										
12am to 6am	0.7494	6.12	0.7494	6.12	0.7494	6.12	0.7494	6.12	0.7494	6.12
Light conditions (base: daylight and dark with artificial light)										
Dawn or dusk	0.1949	5.25	0.1949	5.25	0.1949	5.25	0.1949	5.25	0.1949	5.25
Dark	0.3559	2.79	0.3559	2.79	0.3559	2.79	0.3559	2.79	0.3559	2.79
Weather conditions (base: clear)										
Rain and Snow	-0.1997	-2.98	-0.1997	-2.98	-0.1997	-2.98	-0.1997	-2.98	-0.1997	-2.98
Latent Variables										
Risky behavior: driver vehicle	-0.5581	-20.82	-0.5581	-20.82	-0.0490	-3.03	-0.0490	-3.03	-0.0490	-3.03
Risky behavior: other vehicle	0.0793	3.33	0.0793	3.33	0.5409	11.83	0.5409	11.83	0.5409	11.83
Distracted/careless behavior: driver vehicle	0.5527	17.71	0.5527	17.71	0.5527	17.71	0.5527	17.71	0.5527	17.71
Distracted/careless behavior: other vehicle	1.2623	2.85	1.2623	2.85	1.2623	2.85	1.2623	2.85	1.2623	2.85

Crashes in rain and snow (inclement weather) are less severe in terms of injury across occupants in all seat positions. It is likely that this is a manifestation of the slower speeds and more care exercised by drivers under such environmental conditions. In terms of roadway characteristics, crashes that occur on roadways with a low speed limit of 35 mph or less are generally less severe for passengers in all seat positions. Crashes at non-intersections (access or not a junction, other type of junction) are likely to be more severe for all occupants; this is likely due to higher speeds at non-intersection locations and the lack of traffic control at such locations.

Finally, the two latent variables are found to be very significant in their effects on injury severity (see Table 3). An interesting finding is that risky driving behavior is associated with lower levels of injury severity for all occupants in the driver's vehicle. This finding is actually not that counter-intuitive. Risky drivers may actually be more capable drivers in terms of their agility and ability to swerve and reduce crash severity (Roberti, 2004). The occupants of the vehicle of the non-risky driver who may not be anticipating a crash may therefore be more prone to suffering the more severe outcomes. Moreover, the non-risky drivers are likely to be older and female – and it is possible that these groups are more susceptible to severe injury. Risky driving behavior is associated with greater impact (in terms of injury severity) on the occupants of the other vehicle, which is very much consistent with expectations. Distracted and careless driving behavior is associated with more severe injury outcomes for *both* vehicles. These results illustrate the cross-effects of the behavior of one driver on the injury severity outcomes of occupants in the other vehicle.

6.3 Model Goodness-of-Fit

The performance of the GHDM structure used in this paper can be compared to the one that does not consider latent constructs, maintaining the same specification of the final model. However, this would not constitute a fair specification to test the GHDM specification. Therefore, a model specification that includes the determinants of the latent constructs as explanatory variables, while maintaining the recursivity in the dimensions as obtained from the final GHDM, was estimated. The proof model is an independent model in that the error term correlations across the dimensions are ignored, but the best specification of the explanatory variables (including those used in the GHDM in the structural equation system to explain the latent constructs) is considered to explain the injury severity of the vehicle occupants. The model that has no latent constructs takes the form of a multivariate probit model. This may be referred to as an independent heterogeneous data model (or IHDM). The GHDM and the IHDM specifications are not nested, but they may be compared using the composite likelihood information criterion (CLIC) introduced by Varin and Vidoni (2005). The CLIC takes the following form:

$$\log L^{*}_{CML}(\hat{\boldsymbol{\theta}}) = \log L_{CML}(\hat{\boldsymbol{\theta}}) - tr\left[\hat{\boldsymbol{J}}(\hat{\boldsymbol{\theta}})\hat{\boldsymbol{H}}(\hat{\boldsymbol{\theta}})^{-1}\right] \tag{4}$$

The model that provides a higher value of CLIC is preferred. The performance of the two models may also be compared through the likelihood values $\mathscr{L}(\hat{\boldsymbol{\theta}})$. The corresponding IHDM predictive log-likelihood value may also be computed. The goodness of fit indicators are not presented in Table 3 in the interest of brevity. It was found that the GHDM consistently outperformed the IHDM in every measure of fit, lending credence to the notion that ignoring endogeneity in models of injury severity and driving behavior is likely to yield erroneous predictions of the impacts of safety interventions and engineering designs on crash outcomes. Not only does the GHDM account for endogeneity, but it also offers a flexible methodological framework to measure cross-vehicle driver behavior effects.

7. CONCLUSIONS

This paper presents a comprehensive model of crash injury severity for two-vehicle crashes of all types. The paper employs the GHDM and exploits its methodological capabilities to advance the state of crash severity modeling in three key ways. First, the model system constitutes a simultaneous equations model system capable of accounting for (two) latent driver behavior constructs that influence crash severity outcomes. Second, the model system is able to jointly model the injury severity outcomes for all vehicle occupants in the context of their respective seat positions. Third, the model system accounts for endogeneity in specific explanatory factors such as seat belt use and alcohol involvement. Treating these variables as exogenous variables, when in fact they are endogenous, may lead to inconsistent and biased parameter estimates. Moreover, the model offers the ability to estimate cross-effects, i.e., the effects of the behavior of one vehicle's driver on the injury severity outcomes experienced by occupants in the second vehicle.

It is important to model the injury severity of multiple individuals involved in a crash as different occupants may experience different levels of injury severity. Those differences may be based on observed factors (seat belt use, vehicle type, position of seating) and on unobserved factors (such as the vehicle condition or psychological traits of the driver). Some of these unobserved factors may affect all of the individuals in the same vehicle, while others may impact every person involved in the crash (even across multiple vehicles). The presence of these common unobserved elements motivates the development of a joint multivariate injury-severity model such as that presented in this paper.

It is found that older drivers are particularly susceptible to severe injury outcomes; their impaired driving ability and frail physical condition likely contributes to adverse injury outcomes. Safety interventions inside vehicles and on the roadway should be targeted towards older drivers as their presence in the driving population increases in size. Similarly, interventions that enhance safety at night (such as improved lighting) can help reduce injury severity outcomes. Campaigns that encourage seat belt use and discourage alcohol-impaired driving should be strengthened as these aspects are associated with less severe injury outcomes. Children are safest in the rear seats as they experience less severe injuries when seated there. On the other hand, it is found that passengers in the rear seats suffer more severe injuries in older cars, potentially because many older cars may not have safety features (such as airbags) in the rear. Access control (fewer driveways) on high speed trafficways will improve safety outcomes. Efforts should be made to reduce distracted and careless driving, and vehicular features that may contribute to such driving behavior need to be engineered and designed with care. Distracted and careless driving behavior is associated with worse injury severity for both the driver's vehicle occupants and the other vehicle occupants.

ACKNOWLEDGEMENTS

The authors acknowledge CAPES –**Coordenação de Aperfeiçoamento de Pessoal de Nível Superior** and the Brazilian Government for the funding and support of one of the authors.

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*, *50*, 74-78.

Beanland, V., Sellbom, M., & Johnson, A. K. (2014). Personality domains and traits that predict self-reported aberrant driving behaviours in a southeastern US university sample. *Accident Analysis & Prevention*, *72*, 184-192.

Bhat, C. R. (2015). A new generalized heterogeneous data model (GHDM) to jointly model mixed types of dependent variables. *Transportation Research Part B: Methodological*, *79*, 50-77.

Bilston, L. E., Du, W., & Brown, J. (2010). A matched-cohort analysis of belted front and rear seat occupants in newer and older model vehicles shows that gains in front occupant safety have outpaced gains for rear seat occupants. *Accident Analysis & Prevention*, *42*(6), 1974-1977.

Castro, M., Paleti, R., & Bhat, C. R. (2013). A spatial generalized ordered response model to examine highway crash injury severity. *Accident Analysis & Prevention*, *52*(0), 188-203.

Constantinou, E., Panayiotou, G., Konstantinou, N., Loutsiou-Ladd, A., & Kapardis, A. (2011). Risky and aggressive driving in young adults: Personality matters. *Accident Analysis & Prevention*, *43*(4), 1323-1331.

Donmez, B., & Liu, Z. (2015). Associations of distraction involvement and age with driver injury severities. *Journal of safety research*, *52*, 23-28.

Durbin, D. R., Jermakian, J. S., Kallan, M. J., McCartt, A. T., Arbogast, K. B., Zonfrillo, M. R., & Myers, R. K. (2015). Rear seat safety: Variation in protection by occupant, crash and vehicle characteristics. *Accident Analysis & Prevention*, *80*, 185-192.

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., 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.

Fleiter, J. J., Lennon, A., & Watson, B. (2010). How do other people influence your driving speed? Exploring the 'who'and the 'how'of social influences on speeding from a qualitative perspective. *Transportation research part F: traffic psychology and behaviour*, *13*(1), 49-62.

Ju, Y. H., & Sohn, S. Y. (2011). Quantification method analysis of the relationship between occupant injury and environmental factors in traffic accidents. *Accident Analysis & Prevention*, *43*(1), 342-351.

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.

Mannering, F. L., & Bhat, C. R. (2014). Analytic Methods in Accident Research: Methodological Frontier and Future Directions. *Analytic Methods in Accident Research*, *1*, 1-22.

McEvoy, S. P., Stevenson, M. R., & Woodward, M. (2007). The prevalence of, and factors associated with, serious crashes involving a distracting activity. Accident Analysis & Prevention, 39(3), 475-482.

Mayrose, J., & Priya, A. (2008). The safest seat: effect of seating position on occupant mortality. *Journal of safety research*, *39*(4), 433-436.

National Highway Traffic Safety Administration (NHTSA). (2014). Traffic Safety Facts 2013: A Compilation of Motor Vehicle Crash Data from the Fatality Analysis Reporting System and the General Estimates System, Washington, DC.

Owsley, C., Ball, K., McGwin Jr, G., Sloane, M. E., Roenker, D. L., White, M. F., & Overley, E. T. (1998). Visual processing impairment and risk of motor vehicle crash among older adults. Jama, 279(14), 1083-1088.

Paleti, R., Eluru, N., & Bhat, C. R. (2010). Examining the influence of aggressive driving behavior on driver injury severity in traffic crashes. *Accident Analysis & Prevention*, *42*(6), 1839-1854.

Petridou, E., & Moustaki, M. (2000). Human factors in the causation of road traffic crashes. *European journal of epidemiology*, *16*(9), 819-826.

Roberti, J. W. (2004). A review of behavioral and biological correlates of sensation seeking. *Journal of research in personality*, *38*(3), 256-279.

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.

Stapleton, D.C. (1978). Analyzing political participation data with a MIMIC Model. *Sociological Methodology*, 52-74.

Sümer, N. (2003). Personality and behavioral predictors of traffic accidents: testing a contextual mediated model. *Accident Analysis & Prevention*, *35*(6), 949-964.

Taubman-Ben-Ari, O., & Yehiel, D. (2012). Driving styles and their associations with personality and motivation. *Accident Analysis & Prevention*, *45*, 416-422.

Weiss, H. B., Kaplan, S., & Prato, C. G. (2014). Analysis of factors associated with injury severity in crashes involving young New Zealand drivers. Accident Analysis & Prevention, 65, 142-155.

World Health Organization. (2013). Global status report on road safety 2013. Geneva, Switzerland: World Health Organization.

Yasmin, S., & Eluru, N. (2013). Evaluating alternate discrete outcome frameworks for modeling crash injury severity. *Accident Analysis & Prevention*, *59*, 506-521.

Yasmin, S., Eluru, N., Pinjari, A. R., & Tay, R. (2014). Examining driver injury severity in two vehicle crashes–A copula based approach. *Accident Analysis & Prevention*, *66*, 120-135.

Zhao, N., Reimer, B., Mehler, B., D'Ambrosio, L. A., & Coughlin, J. F. (2013). Self-reported and observed risky driving behaviors among frequent and infrequent cell phone users. *Accident Analysis & Prevention*, *61*, 71-77.