Multi-vehicle Collisions involving Large Trucks on Highways: An Exploratory Discrete Outcome Analysis

Mouyid Islam, E.I.T.*
Graduate Research Assistant - Department of Civil Engineering
(915) 345-7226, mbislam@miners.utep.edu

Salvador Hernandez, PhD
Assistant Professor - Department of Civil Engineering
(915) 747-6348, shernandez43@utep.edu

500 West University Avenue
University of Texas at El Paso
El Paso, TX 79968

Abstract:

Trucking industry is considered a driving force for logistic and supply chain systems which indirectly influences the national economy. So, any impedance in truck-flow or supply chain system eventually brings substantial consequences in terms of monetary values. As such, a growing concern related to large-truck (Gross Vehicle Weight Rating (GVWR) greater than 10,000 pounds) crashes has increased in recent years due to the potential economic impacts and level of injury severity sustained. With this in mind, this study aims to analyze the injury severities of multi-vehicle collisions involving large-trucks through an advanced econometric modeling approach to shed light on the contributing factors leading to large-truck crashes. Through a fused national crash datasets, we hope to provide a clearer understanding of the complex interactions of contributing factors (e.g., factors related to human (drivers), vehicle, and road-environment) influencing multi-vehicle crash outcomes. To capture these complexities using the national crash database and understand the underlying causal factor, discrete outcome models namely random parameter ordered probit and mixed logit (which accounts for observable factors) were estimated to predict the likelihood of five injury severity outcomes—fatal, incapacitating, non-incapacitating, possible injury, and no injury. Estimation findings indicate that the level of injury severity is highly influenced by a number of complex interactions of factors and that the effect of the some of the factors can vary across the observations.

Keywords: multi-vehicle, large truck, mixed logit, ordered probit, injury severity

corresponding author

INTRODUCTION

Trucking is a vital lifeline to any growing economy and is considered as the driving force for economy primarily based on logistic and supply chain systems. Hence, any disruption in the trucking logistics system clearly indicates substantial ramifications. As such, a growing concern related to large-truck (Gross Vehicle Weight Rating (GVWR) greater than 10,000 pounds) crashes has increased in recent years due to the potential economic impacts and sustained injury severities. Recent statistical data indicate that large trucks account for more fatalities in the United States (US) than passenger vehicles based on the number of registered vehicles and vehicle-miles traveled (VMT) (FHWA, 2010; NHTSA, 2008). Although large trucks accounted for four percent of registered vehicles and eight percent of VMT in 2008, eleven percent of motor vehicle crash deaths in 2008 were a result of large truck involvement in the crashes (FHWA, 2010). Although large trucks heavily impact the national economy through daily freight movements, the level of injury severity of the collision partners are also high coupled with incurring high societal cost associated with fatalities, injuries and property damages.

The cost associated with these large-truck involved crashes is of great concern and obviously substantial as evidenced from the fact the estimated cost of a police-reported crash involving a large truck on average $91,112 based on 2005 dollars (Zaloshnja and Miller, 2006). In addition, the study also estimated the average cost per fatality, non-fatality and property-damage-only as $3,604,518, $195,258, and $15,114, respectively (Zaloshnja and Miller, 2006). An earlier study by Zaloshnja and Miller (2004) estimated the cost associated with multiple combination trucks as having the highest cost of $88,483 per crash based on 2000 dollars. The crash costs based on 2000 dollars per 1000 truck miles were $157 for single unit trucks, $131 for single combination trucks, and $63 for multiple combinations (Zaloshnja and Miller, 2004). The above mentioned costs illustrate the potential monetary impacts on society. Therefore, any increase in the number and level of crash severity is of great concern to transportation organizations that operate, maintain, and construct the transportation system as well as to trucking companies.

Considering all the crashes involving large trucks and associated societal costs, this study aims to analyze the injury severities of multi-vehicle large-truck involved crashes through an advanced econometric modeling approach namely mixed logit model to shed light on the contributing factors leading to crashes involving large truck. To achieve this, the three datasets from the National Automotive Sampling System General Estimated System (NASS-GES) were fused to obtain a crash sample. Utilizing this fused dataset, this study is intended to provide an improved understanding of the complex interactions of contributing factors (e.g., factors related to individuals (drivers), vehicle, and road-environment) influencing large-truck crash outcomes through this fused dataset. To capture these complexities using the above mentioned database, consideration of random parameters provides a mechanism to account for any unobserved heterogeneity that may exist—that is, unobserved factors that may vary across observations. To the best of authors' knowledge these are the first attempts at modeling multi-vehicle large truck severity utilizing NASS-GES dataset. Although the mixed logit model have been applied to large-truck crash severity (see literature review) from different modeling perspectives, this research extends the current literature and introduces additional significant variables related to human factors in regards to multi-vehicle large-truck crashes (also a first).

METHODOLOGY

In order to achieve a better understanding of the injury severity of large truck involved multi-vehicle crashes with discrete outcome models, we developed random parameter ordered probit and mixed logit models.

Ordered Porbit Framework: We developed a random parameter ordered probit model to capture the injury severity experienced while accounting for unobserved heterogeneity (McKelvey and Zavoina, 1975; Chistoforou et al., 2010; Zhu and Srinivasan, 2011) because of the ordinal nature of injury according to the KABCO scale (i.e., ‘K’ for Fatal, ‘A’ for Incapacitating injury, ‘B’ for Non-incapacitating Injury, ‘C’ for Possible Injury and ‘O’ for Property-Damage-Only). In this study, we followed the descending order (i.e., 0 for K , 1 for A, 2 for B, 3 for C and 4 for O) (Islam and Hernandez, 2012) rather than ascending order in the previous studies (Chistoforou et al., 2010; Abdel-Aty, 2003; Gray et al., 2008; Kockelman and Kweon, 2002; Lee and Abdel-Aty, 2005; O’Donnell and Connor, 1996; Pai and Saleh, 2008; Quddus et al., 2002; Xie et al., 2009; Zajac and Ivan, 2002) to account for any bias resulting from under-reporting tendency in the crash and variability of parameter estimation (Ye and Lord, 2011).

In the formulation of the model, an unobserved variable y^* is a modeling basis of ordinal ranking of the data, with y^* specified as a latent and continuous measure of injury severity of each observation (Washington et al., 2010):

$$y^* = \boldsymbol{\beta X} + \varepsilon \tag{1}$$

where:

- y^* : is the dependent variable (specified as a latent and continuous measure of injury severity of each observation *n*),
- $\boldsymbol{\beta}$: is a vector of estimable parameters,
- $\boldsymbol{X}$: is a vector of explanatory variables (e.g., human, roadway segment, vehicle, and crash mechanism characteristics),
- ε : is a random error term (assumed to be normally distributed with 0 mean and a variance of 1).

Using Equation 1, and under the order probit framework the observed ordinal data *y* (e.g., injury severity) for each observation can be represented as (Washington et al., 2010):

$$\begin{aligned}
&y = 0 \quad if\ y^* \le \mu_0 \\
&y = 1 \quad if\ \mu_0 \le y^* \le \mu_1 \\
&y = 2 \quad if\ \mu_1 \le y^* \le \mu_2 \\
&y = \cdots \\
&y = I \quad if\ y^* \ge \mu_{I-2}
\end{aligned} \tag{2}$$

where:

- μ : are estimable parameters (i.e., thresholds) that define y and are estimated jointly with the model parameters $\boldsymbol{\beta}$, which corresponds to integer ordering, and I is the highest integer ordered response (e.g., PDO this is 4).

To estimate the probabilities of I specific ordered response for each observation n, ε is assumed to be normally distributed with 0 mean and variance. Hence, the ordered probit model with ordered selection probabilities is defined as follows:

$$
\begin{aligned}
&P_n(y=0) = \Phi(-\boldsymbol{\beta X}) \\
&P_n(y=1) = \Phi(\mu_1 - \boldsymbol{\beta X}) - \Phi(-\boldsymbol{\beta X}) \\
&P_n(y=2) = \Phi(\mu_2 - \boldsymbol{\beta X}) - \Phi(\mu_1 - \boldsymbol{\beta X}) \\
&\ldots \\
&P_n(y=I) = 1 - \Phi(\mu_{I-2} - \boldsymbol{\beta X})
\end{aligned} \tag{3}
$$

where:

$P_n(y=I)$: is the probability that observation n has I as the highest ordered-response index (in our case PDO being 4 is the highest)

$\Phi(\cdot)$: is the standard normal cumulative distribution function

Marginal effects are computed at the sample mean for each category (Greene, 2007; Washington et al., 2010):

$$\frac{P_n(y=I)}{\partial \boldsymbol{X}} = [\phi(\mu_{I-1} - \boldsymbol{\beta X}) - \phi(\mu_I - \boldsymbol{\beta X})]\boldsymbol{\beta} \tag{4}$$

where:

$\phi(\cdot)$: is the probability mass function of the standard normal distribution

Greene (2007) developed an estimation procedure that utilizes simulated maximum likelihood estimation to incorporate random parameters in the ordered probit modeling scheme. The random parameter ordered probit model is formulated by taking into account of an error term being correlated with the unobserved factors in ε_i (as shown in Equation 1) which translates the individual heterogeneity into parameter heterogeneity as follows (Greene, 2007):

$$\beta_{in} = \beta + \gamma_{in} \tag{5}$$

where:

γ_{in} : is randomly distributed term (for example a normally distributed term with mean 0 and variance σ^2).

Mixed Logit Framework: In terms of utility functions and other methodological flexibility, we seek to develop mixed logit model that can be used to determine the contributing factors that influence the likelihood of severity outcomes in large truck involved crashes.

S_{in} as linear function that determines discrete outcome i as injury severity outcome such as fatality, incapacitating injury, non-incapacitating injury, possible injury, and no-injury (Property-Damage-only) for observation n such that: (Washington et al., 2010):

$$S_{in} = \mathbf{\beta}_i \mathbf{X}_{in} + \varepsilon_{in} \tag{6}$$

where,

$\boldsymbol{X_{in}}$: is vector of explanatory variables covering driver, vehicle road and environment factors that determine injury outcome (i),

$\boldsymbol{\beta}_{\mathrm{i}}$: is vector of estimable parameters,

ε_{in} : is random error.

If ε_{in}'s are assumed to be generalized extreme value distributed, McFadden (1981) has shown that the multinomial logit results such that:

$$P_n(i) = \frac{EXP[\boldsymbol{\beta}_i \mathbf{X}_{in}]}{\sum_I EXP[\boldsymbol{\beta}_i \mathbf{X}_{in}]} \tag{7}$$

where, $P_{\mathrm{n}}(i)$ is probability of observation n having severity outcome i (such as fatality, incapacitating injury, non-incapacitating injury, possible injury, PDO) ($i \in I$ with I denoting all possible outcomes of injury severity for observation n).

As NASS-GES crash data are likely to have a significant amount of unobserved heterogeneity because of the investigating police discretion and their reported estimates of the representative crash data sample all over the USA (for instance, the factors influencing the severity outcome more or less severe or no-injury), we consider the possibility that elements of the parameter vector β_i may vary across observations of each crash by using a random-parameters logit model (also known as the mixed logit model). Previous work by McFadden and Rudd (1994), Geweke et al. (1994), Revelt and Train (1997, 1999), Train (1997), Stern (1997), Brownstone and Train (1999), McFadden and Train (2000), and Bhat (2001) have shown the development and effectiveness of the mixed logit approach which can explicitly account for the variations (across crash observations) of the effects that variables have on the severity outcomes (or choices) considered in this study. The mixed logit model is written as (see Train, 2003),

$$P_n(i) = \int \frac{EXP[\boldsymbol{\beta}_i \mathbf{X}_{in}]}{\sum_I EXP[\boldsymbol{\beta}_i \mathbf{X}_{in}]} f(\boldsymbol{\beta}_i|\boldsymbol{\varphi}) d\boldsymbol{\beta}_i \tag{8}$$

where, $f(\boldsymbol{\beta}_i|\varphi)$is the density function of $\boldsymbol{\beta}_i$, $\boldsymbol{\varphi}$ is a vector of parameters of the density function (mean and variance), and all other terms are as previously defined. This model can now account for severity outcome specific variations of the effect of $\mathbf{X}_{in}$ on severity outcome probabilities, with the density function $f(\boldsymbol{\beta}_i|\boldsymbol{\varphi})$ used to determine $\boldsymbol{\beta}_i$. Mixed logit probabilities are then a weighted average for different values of $\boldsymbol{\beta}_i$ across crash observations where some elements of the vector $\boldsymbol{\beta}_i$ may be fixed and some randomly distributed. If the parameters are random, the mixed logit weights are determined by the density function $f(\boldsymbol{\beta}_i|\boldsymbol{\varphi})$ (Milton et al., 2008; Washington et al., 2010).

In order to estimate the impact of particular variables on the injury-outcome likelihood, elasticities (or direct-pseudo elasticity) are computed. In the context of the current injury severity model, most of the variables are indicator in nature, direct-pseudo elasticities are estimated to measure the marginal effects of indicator variables when any particular indicator variable switches from 0 to 1 or reverse (Washington et al., 2010). Also, this is translated to percentage change in the likelihood of while the indicator variables switching between 0 and 1 or 1 to 0. For binary indicator variables, the direct-pseudo elasticity is estimated as follows (Kim et al., 2010):

$$E_{x_{nk}(i)}^{P_n(i)} = \frac{P_n(i)[given\ x_{nk}(i)=1] - P_n(i)[given\ x_{nk}(i)=0]}{P_n(i)[given\ x_{nk}(i)=0]} \tag{9}$$

where, $P_n(i)$ is given the Equation (8) and simulated as shown in Equation (10).
x_{nk} (i) = the k-th independent variable associated with injury severity i for observation n.

The unconditional probability in Equation (8) (Kim et al., 2010) can be estimated with an unbiased and smooth simulator (McFadden and Train, 2000) that is computed as (Walker and Ben-Akiva, 2002; Kim et al., 2010):

$$\hat{P}_n(i) = \frac{1}{R}\sum_{r=1}^{R} P_{in} = \frac{1}{R}\sum_{r=1}^{R} \int \frac{EXP[\boldsymbol{\beta}_i \mathbf{X}_{in}]}{\sum_I [\boldsymbol{\beta}_i \mathbf{X}_{in}]} f(\boldsymbol{\beta}_i|\boldsymbol{\varphi}) d\boldsymbol{\beta}_i \tag{10}$$

where, R = the total number of draws.

Since the direct pseudo-elasticity is calculated for each observation, it is usually reported as the average direct pseudo-elasticity (taking average over the sample) as a measure of the marginal effect of an indicator variable on the likelihood of a particular injury severity outcome (Kim et al., 2010).

With the simulator in Equation (10), Maximum Simulated Likelihood Estimation (MSLE) can be used to estimate parameters and this MSLE estimator is asymptotically normal and consistent (Lee, 1992; Kim et al., 2010):

$$max_{\boldsymbol{\beta}_{in}} \sum_{n=1}^{N} \sum_{i=1}^{I} y_{in} ln\hat{P}_n(i) \tag{11}$$

where, N = the total number of observations (i.e., crashes in the sample)
y_{in} = 1 if individual n suffers from injury severity i, 0 otherwise.

Maximum likelihood estimation with random parameters of both models – mixed logit and random parameter ordered probit models is undertaken with simulation approaches due to the difficulty in computing the probabilities (Halton, 1960; Train, 1999; Bhat, 2003; Milton et al., 2008; Anastasopoulos and Mannering, 2009). The most widely accepted simulation approach utilizes Halton draws which is a technique developed by Halton (1960) to generate a systematic non-random sequence of numbers. Halton draws have been shown to provide a more efficient distribution of the draws for numerical integration than purely random draws (Bhat, 2003; Train, 1999, Christoforou et al., 2010).

EMPIRICAL SETTINGS

The data for large trucks involved crashes was obtained from the nationwide NASS-GES crash database maintained by National Highway Traffic Safety Administration (NHTSA). A large truck is commonly classified as a tractor-trailer, single-unit truck, or cargo van having GVWR greater than 10,000 pounds (IIHS, 2009). The GES database is based on a nationally representative probability sample selected from the estimated 6.4 million police-reported crashes resulting in a fatality or injury and those involving major property damage annually (NASS-GESS, 2005). To investigate contributing factors on human, vehicle, and road-environment, a sample of 8,291 data observations were extracted from this large dataset representing crashes only involving large trucks for the interstate highway system over a period of four years, from 2005 to 2008. The maximum level of injury severity recorded in the vehicle or person dataset was aggregated to represent a crash. So, each observation in the sample is a crash representing the maximum level of injury of the occupants, involving at least one large truck with one or more vehicles on interstate highways. The crash dataset was fused to vehicle and person dataset through appropriate linking variable, the crash number; while vehicle and person dataset were linked through vehicle and crash number using the Statistical Analysis System (SAS). The mixed logit and ordered probit frameworks were modeled in Limdep (NLOGIT 4.0).

Table 1 and Table 2 show the descriptive statistics of key variables in the models. Although some of the variables are common in both models, the data description of some important variables is presented here. With regards to ordered probit model, Table 1 illustrates about 33% observations related to side-swipe in the same directions, 81% related to rollover collisions resulted in multi-vehicle collisions. Additionally, as seen from Table 1, lane changing maneuvers account for 12% of the total observations compared to 65.2% regarding going straight. Another key observation to note from the data is that dark condition and summer months (i.e., June to August) accounting for 11% and 23.5% in multi-vehicle collisions, respectively. The statistics further illustrate that speeding and struck by other vehicles account for about 8% and 46.6% of the total observations for fatalities multi-vehicle collisions, respectively.

Table 1 Descriptive Statistics of Key Variables in Ordered Probit Model

Meaning of Variables in the Model	Mean	Std. Dev.
Vehicle maneuver during pre-crash situation (1 if left or right side departure, 0 otherwise)	0.041	0.199
Light condition of street (1 if dark, 0 otherwise)	0.109	0.312
Passenger role (1 if passenger is present, 0 otherwise)	0.977	0.146
Vehicle maneuver during pre-crash situation (1 if going straight, 0 otherwise)	0.652	0.476
Driver's attention level at the time of impending crash (1 if distraction or in attention, 0 otherwise)	0.041	0.197
Role as crash partner (1 if struck, 0 otherwise)	0.466	0.498
The most harmful event in crash consequences (1 if rollover, 0 otherwise)	0.809	0.392
Orientation of vehicle at the time of crash (1 if sideswipe in the same direction, 0 otherwise)	0.331	0.471
Occupants' use of available vehicle restraints (1 if no restraint used, 0 otherwise)	0.018	0.133
Months of year (1 if summer months (June to August), 0 otherwise)	0.235	0.424
Gender of the occupants (1 if male, 0 otherwise)	0.938	0.240
Drivers' working/residing place according to license record (1 if Texas, 0 otherwise)	0.100	0.300
Speed-related factor in crash (1 if speed as a factor, 0 otherwise)	0.079	0.270
Vehicle maneuver just prior to impending crash (1 if changing lane, 0 otherwise)	0.118	0.323
Trailing unit when the crash occurred (1 if one trailer, 0 otherwise)	0.750	0.432

Table 2 shows that about 42.4% of the total crash observations related to rear-end collisions and on average more than two (i.e., 2.3 vehicles) vehicles involved in multi-vehicle collisions. The statistics as seen in Table 2 illustrate that lane changing maneuver, inattentive driving, dark condition accounts for 11.8%, 4.1% and 11% of the total crash observations, respectively. Curved sections of highways and wet pavement account for 8.1% and 15.2% of total crash observations, respectively. The time specific variables such as summer month (i.e., June to August) and time of day – 2 pm and 5 am on average account for 23.5%, 5.5%, and 12.3% of total crash observations, respectively.

Table 2 Descriptive Statistics of Key Variables in Mixed Logit Model

Meaning of Variables in the Model	Mean	Std. Dev.	Outcome
Vehicle maneuver during pre-crash situation (1 if left or right side departure, 0 otherwise)	0.009	0.098	Fatal
Light condition of street (1 if dark, 0 otherwise)	0.109	0.312	
Orientation of vehicle at the time of crash (1 if head-on, 0 otherwise)	0.008	0.093	
Time of the day (1 if 2 pm in the afternoon, 0 otherwise)	0.055	0.228	
Driver's attention level at the time of impending crash (1 if distraction or in attention, 0 otherwise)	0.041	0.197	Incapacitating Injury
Time of the day (1 if 5 am in the morning, 0 otherwise)	0.123	0.328	
Vehicular factors (1 if tire-related malfunction, 0 otherwise)	0.007	0.085	
Orientation of vehicle at the time of crash (1 if rear-end, 0 otherwise)	0.424	0.494	
The most harmful event in crash consequences (1 if rollover, 0 otherwise)	0.809	0.392	
Orientation of vehicle at the time of crash (1 if sideswipe in the same direction, 0 otherwise)	0.331	0.470	Non-Incapacitating Injury
Occupants' use of available vehicle restraints (1 if no restraint used, 0 otherwise)	0.018	0.133	
Time of the day (1 if 4 am in the morning, 0 otherwise)	0.020	0.141	
Months of year (1 if summer months (June to August), 0 otherwise)	0.235	0.424	
Gender of the occupants (1 if male, 0 otherwise)	0.938	0.240	Possible Injury
Drivers' working/residing place according to license record (1 if Texas, 0 otherwise)	0.100	0.300	
Number of vehicles involved in the crash	2.324	0.672	
Speed-related factor in crash (1 if speed as a factor, 0 otherwise)	0.079	0.270	
Road surface condition (1 if wet, 0 otherwise)	0.152	0.359	
Vehicle maneuver just prior to impending crash (1 if changing lane, 0 otherwise)	0.118	0.323	No-injury (PDO)
Light condition of street (1 if the surrounding area is dark but outside is lighted, 0 otherwise)	0.157	0.364	
Driver's attention level at the time of impending crash (1 if sleepy, 0 otherwise)	0.002	0.042	
Trailing unit when the crash occurred (1 if one trailer, 0 otherwise)	0.751	0.432	
Orientation of vehicle at the time of crash (1 if rear-end, 0 otherwise)	0.424	0.494	
Alignment of highway section (1 for curved section, 0 otherwise)	0.081	0.274	

The correlation matrix for both of the severity models – random parameter ordered probit and mixed logit was computed. The correlation matrix for random parameter ordered probit model indicate that lane changing maneuver has correlation coefficients of 0.501 and 0.329 with going straight and side-swipe collisions, respectively. On the other hand, the correlation matrix for mixed logit model indicate that rear-end collision has correlation coefficients of 0.604 with side-swipe collisions and time – 4 o'clock hour has correlation coefficient of 0.385 with 5 o'clock hour. Although the magnitude of coefficients might pose some multicollinearity issues, the maneuver and collisions are not seriously correlated in the models. For ordered probit model, lane changing maneuver might result in subsequent actions of going straight and side-swipe in the same direction in multi-vehicle collisions. Similar is true for mixed logit model where rear-end collision might be outcome of some subsequent actions of side-swipe collisions. Also, 4 am and 5 am in the morning indicates early morning hours from 4 to 5 o'clock in the morning account for severe injuries for multi-vehicle collisions.

EMPIRICAL RESULTS

The variables in both estimated models were found to be statistically significant within a 95% and 90% confidence level for random parameter ordered probit and mixed logit models, respectively.

A random parameter ordered probit and mixed logit model was developed based on fixed parameter ordered probit and initial multinomial logit model, respectively. However, random parameter ordered probit model and mixed logit were found statistically superior to base models (i.e., fixed parameter ordered probit model and multinomial logit model) as evidenced from the following hypothesis and likelihood ratio test.

$$\chi^2 = -2[LL_{FIX}(\beta^{FIX}) - LL_{RAN}(\beta^{RAN})] \tag{12}$$

where:

$LL_{FIX}(\beta^{FIX})$: is the log-likelihood at convergence of the fixed parameters model (-3032.560)

$LL_{RAN}(\beta^{RAN})$: is the log-likelihood at convergence of the random parameters model (-3022.542)

χ^2 = 20.036 (5 degree of freedom)

The Chi-square statistic for the likelihood ratio test with five degrees of freedom gave a value greater than the 99.88% ($\chi^2 = 20.036$) confidence interval, indicating that the random parameter model is statistically superior to the corresponding fixed parameter models.

$$\chi^2 = -2[LL_{MXL}(\beta^{MXL}) - LL_{MNL}(\beta^{MNL})] \tag{13}$$

where:

$LL_{FIX}(\beta^{FIX})$: is the log-likelihood at convergence of the multinomial logit model (-3087.115)

$LL_{RAN}(\beta^{RAN})$: is the log-likelihood at convergence of the mixed logit model (-3081.050)

χ^2 = 12.13 (with 3 degree of freedom)

The Chi-square statistic for the likelihood ratio test with three degrees of freedom gave a value greater than the 99.31% ($\chi^2 = 12.13$) confidence interval, indicating that the random parameter model is statistically superior to the corresponding fixed parameter model (i.e., multinomial model). In both cases above, this means that the null hypothesis that the random parameter estimated models (i.e., mixed logit and random parameter ordered probit) are no better than the fixed models (i.e., multinomial and ordered probit model) is rejected.

The contributing factors such as human, road and environment, vehicle and crash mechanisms in the multi-vehicle large truck involved crashes are described below as found in the model results shown in Table 3 and Table 4.

There are five parameters found to be random in the random parameter ordered probit model. These five random parameters are constant, dark condition, side-swipe collision (same direction), lane changing maneuver, and being male occupants. The first parameter – constant, having mean of 6.088 and standard deviation of 2.672, has 4.87% observations below zero (i.e., 91.13% above zero). This captures significant unobserved heterogeneity present in data. The second parameter – dark condition, having mean of -0.269 and standard deviation of 2.223, has 54.82% observations below zero (i.e., 45.18% above zero). This indicates that 54.8% multi-vehicle large truck collisions in the dark condition resulted in severe injuries. The third parameter – side-swipe collision (same direction), having mean of 1.251 and standard deviation of 1.004, has 10.64% of observations below zero (i.e., 89.36% above zero). This indicates that 89.4% multi-vehicle large truck collision as side-swipe (same direction) resulted in less severe injuries. The fourth parameter – lane changing maneuver, having mean of 2.617 and standard deviation of 3.119, has 20.1% observations below zero (i.e., 79.9% above zero). This indicates that 79.9% multi-vehicle large truck collisions as consequences of lane changing maneuver resulted in less severe injuries. The fifth parameter – male occupants, having mean of 0.719 and standard deviation of 0.546, has 9.4% observations below zero (i.e., 89.6% above zero). This indicates that 89.6% of multi-vehicle large truck collisions involving male occupants experienced less severe injuries. The estimated model results are presented in Table 3.

Table 3 Multi-vehicle Random Parameter Ordered Probit Model Results

Injury Severity – Random Parameter Ordered Probit	**Random Parameters Model**		
	Coeff.	***t*-stat**	***P*-value**
Constant	6.088	18.889	0.000
Standard Deviation of parameter distribution	*3.672*	*34.078*	*0.000*
Weather condition (1 if snow, 0 otherwise)	0.861	4.344	0.000
Months of the year (1 if summer months (June - August), 0 otherwise)	-0.580	-7.484	0.000
Light condition of street (1 if dark, 0 otherwise)	-0.269	-2.137	0.033
Standard Deviation of parameter distribution	*2.223*	*17.636*	*0.000*
Trailing unit when the crash occurred (1 if one trailer, 0 otherwise)	1.402	17.427	0.000
Vehicle role (1 if struck by other vehicle, 0 otherwise)	1.522	17.311	0.000
The most harmful event (1 if rollover, 0 otherwise)	1.691	19.231	0.000
Orientation of vehicle at the time of crash (1 if sideswipe in the same direction, 0 otherwise)	1.251	12.857	0.000
Standard Deviation of parameter distribution	*1.004*	*12.524*	*0.000*
Vehicle maneuver just prior to impending crash (1 if changing lane, 0 otherwise)	2.617	11.868	0.000
Standard Deviation of parameter distribution	*3.119*	*18.117*	*0.000*
Vehicle maneuver just prior to impending crash (1 if going straight, 0 otherwise)	0.457	5.427	0.000
Factor of crash identified in the investigation (1 if speed, 0 otherwise)	-0.846	-7.868	0.000
Driver’s attention level at the time of impending crash (1 if distraction or in attention, 0 otherwise)	-1.158	-7.733	0.000
Occupants’ use of available vehicle restraints (1 if no restraint used, 0 otherwise)	-3.250	-17.810	0.000
Location of the occupants in the vehicle (1 if for passenger position, 0 otherwise)	1.018	5.533	0.000
Gender of the occupants (1 if male, 0 otherwise)	0.719	5.598	0.000
Standard Deviation of parameter distribution	*0.546*	*13.735*	*0.000*
Drivers’ working/residing place according to license record (1 if Texas, 0 otherwise)	-0.789	-7.864	0.000
Threshold 1, $\mu 1$	2.845	14.047	0.000
Threshold 2, $\mu 2$	4.708	21.307	0.000
Threshold 3, $\mu 3$	6.280	25.969	0.000
Log-likelihood at zero, LL(0)	-3258.341		
Log-likelihood at convergence, LL(β)	-3022.542		
Chi-squared value	471.598		
McFadden’s pseudo R^2	0.072		
Number of observations, N	6,588		

Since no-injury (i.e., PDO) is a base condition in the mixed logit model, the estimated results presented in Table 4 is the difference between the target injury outcomes (i.e., Fatal, Incapacitating, Non-incapacitating injury outcome) with respect to base condition (i.e., PDO). There are three random parameters found statistically significant for multi-vehicle mixed logit model. Constant specific to fatality, having mean of -8.729 and standard deviation of 2.663, has 99.95% observations below zero. This captures some unobserved heterogeneity present in the fatal outcome in multi-vehicle large truck involved crashes.

Table 4 Multi-vehicle Mixed Logit Model Results

Injury Severity - Mixed Logit	Random Parameters Model		
	Coeff.	***t*-stat**	***P*-value**
Fatal Outcome			
Constant	-8.729	-4.047	0.000
Standard Deviation of parameter distribution	*2.663*	*2.618*	*0.009*
Vehicle maneuver during pre-crash situation (1 if left or right side departure, 0 otherwise)	2.939	2.272	0.023
Light condition of street (1 if dark, 0 otherwise)	2.065	3.298	0.001
Orientation of vehicle at the time of crash (1 if head-on, 0 otherwise)	2.804	2.278	0.023
Time of the day (1 if 2 pm in the afternoon, 0 otherwise)	2.224	3.066	0.002
Incapacitating Injury Outcome			
Constant	-3.027	-14.942	0.000
Driver's attention level at the time of impending crash (1 if distraction or in attention, 0 otherwise)	1.279	2.550	0.018
Vehicular factors (1 if tire-related malfunction, 0 otherwise)	2.276	2.927	0.003
The most harmful event in crash consequences (1 if rollover, 0 otherwise)	-3.233	-1.956	0.051
Standard Deviation of parameter distribution	*2.195*	*2.126*	*0.033*
Orientation of vehicle at the time of crash (1 if rear-end, 0 otherwise)	0.495	2.159	0.031
Time of the day (1 if 5 am in the morning, 0 otherwise)	1.235	4.381	0.000
Non-incapacitating Injury Outcome			
Constant*	-8.233	-2.176	0.029
Standard Deviation of parameter distribution	*4.522*	*1.993*	*0.046*
Occupants' use of available vehicle restraints (1 if no restraint used, 0 otherwise)	4.320	2.219	0.026
Time of the day (1 if 4 am in the morning, 0 otherwise)	2.119	1.831	0.067
Months of year (1 if summer months (June to August), 0 otherwise)	0.852	1.756	0.079
Orientation of vehicle at the time of crash (1 if sideswipe in the same direction, 0 otherwise)	-1.396	-1.816	0.069
Possible Injury Outcome			
Constant	-2.678	-10.091	0.000
Gender of the occupants (1 if male, 0 otherwise)	-0.455	-2.311	0.021
Drivers' working/residing place according to license record (1 if Texas, 0 otherwise)	0.790	5.510	0.000
Speed-related factor in crash (1 if speed as a factor, 0 otherwise)	0.346	1.994	0.046
Number of vehicles involved in the crash	0.245	3.767	0.000
Road surface condition (1 if wet, 0 otherwise)	0.504	3.696	0.037
Non-Injury Outcome (Property-Damage-Only)			
Alignment of highway section (1 for curved section, 0 otherwise)	-0.339	-2.162	0.031
Orientation of vehicle at the time of crash (1 if rear-end, 0 otherwise)	-0.232	-2.161	0.031
Light condition of street (1 if the surrounding area is dark but outside is lighted, 0 otherwise)	0.512	2.940	0.003
Vehicle maneuver just prior to impending crash (1 if changing lane, 0 otherwise)	0.371	2.086	0.037
Driver's attention level at the time of pre-crash (1 if sleepy, 0 otherwise)	-2.188	-3.357	0.001
Trailing unit when the crash occurred (1 if one trailer, 0 otherwise)	0.741	7.510	0.000
Log-likelihood at zero, LL(0)	-10602.98		
Log-likelihood at convergence, , LL(β)	-3081.050		
Chi-squared value	15043.85		
McFadden pseudo-R^2	0.709		
Number of observations, N	6,588		

The second parameter – rollover, having mean of -3.233 and standard deviation of 2.195, has 92.9% observation below zero. This fact indicates that 92.6% multi-vehicles collision associated with rollover resulted in decrease in incapacitating injuries. The third parameter – constant specific to non-incapacitating injury, having -8.233 and standard deviation of 4.522, has 96.6% observation below zero. This captures some unobserved heterogeneity present in the non-incapacitating injury in the multi-vehicle large truck involved collisions.

The comparative statistical performance of both discrete choice models are presented in Table 5. The reported pseudo-R^2 is 0.709 for mixed logit model, in contrast to 0.072 in random parameter ordered probit model, which implying the mixed logit model fits the data predicting the multi-vehicle collisions for all five injury outcomes. It is clearly found that log-likelihood at convergence is much better for mixed logit model than random parameter ordered probit model as well as their corresponding Chi-squared value. Table 5 shows number of observations of mixed logit model is five times (5x6588) of the original observations (i.e., 6588) because of number of outcomes (i.e., Fatal, Incapacitating, Non-incapacitating, possible injury and No-injury) considered in the modeling framework in Limdep software. Also, the number of parameters for mixed logit (means and standard deviation of parameter distribution) and random parameters ordered probit (means and standard deviation of parameter distribution as well as the thresholds in the framework) models are software specific parameter reporting.

Table 5 Comparison of Mixed Logit with Random Parameter Ordered Probit Model

Methodological Evaluation	Mixed Logit Model	Random Parameter Ordered Probit Model
Number of observations in the model framework	32,940 (5x6,588)	6,588
Restricted log-likelihood	-10602.980	-3258.341
Log-likelihood at convergence	-3081.050	-3022.542
Chi-squared value	15043.85	471.598
McFadden Pseudo R^2	0.709	0.072
Number of random parameters	3	5
Number of parameters	31	24

The parameter estimates were not only measure of the variables; marginal effects in terms of average direct pseudo-elasticities were also computed to measure the impact of respective variables for mixed logit model on the corresponding injury outcome. The average direct pseudo-elasticities are presented in Table 6.

Human Factors: Male occupants are 38.3% less likely to be involved in possible injuries as supported by a study by Chen and Chen (2011) for fatal or incapacitating and non-incapacitating or possible injuries. Distracted driving is 7.7% more likely to result in incapacitating injuries because of multiple vehicular interaction dynamics. Not wearing seat-belt by the occupants result in 11.5% more likelihood of involved with non-incapacitating injuries which might indicate the unbelted occupants rather than drivers. Similar findings by Chen and Chen (2011) also indicated the fact is true for fatal or incapacitating and non-incapacitating or possible injuries. Drivers residing or registered to work in the state of Texas are 12.8% more likely to be involved with possible injuries. Sleepy driving condition is more likely to be not involved with non-injury which indirectly shows the more likely to be involved with serious injuries. As found in the random parameter ordered probit model, presence of passenger reduces the likelihood of severe injuries which might indicate the passenger keeps awake the drivers on their driving task.

Road and Environment Factors: Dark condition leads to 50.3% more likelihood of fatalities since other vehicles might be completely blinded by such unfavorable driving conditions. This fact is supported by a similar study by Chen and Chen (2011) where dark light condition increases both – fatal or incapacitating injuries and non-incapacitating or possible injuries. Similarly, dark but lighted condition increases possible injuries by 9.4%. Time of the day and month of the year implies the traffic condition on the highway. Time of day such as 2 o' clock in the afternoon increases 30.1% likelihood of fatalities which clearly indicates drowsy driving after-lunch effects by other vehicles (other than trucks). Similarly, 4 and 5 o'clock in the morning also increases 3.5% likelihood of non-incapacitating and 23.9% incapacitating injuries, respectively which indicates of sleepy or drowsy driving condition. Summer months (from June to August) increases the 11% likelihood of non-incapacitating injuries because of more traffic on the highways and great chances of interaction of vehicles leading to collisions. Whereas, the wet pavement condition increases 9.4% likelihood of possible injuries because of unfavorable driving and braking on slippery road condition for other vehicles and also braking characteristics of large trucks. Curved segments of the highways decrease the likelihood of non-injury crashes which indirectly points towards serious injuries. Random parameter ordered probit model indicated that snow reduces the likelihood of severe injuries as drivers are very cautious maneuvering the vehicles in the adverse weather condition.

Vehicular Factors: Tire related malfunction increases 5.2% likelihood of incapacitating injury which indicates the ignorance of vehicle maintenance of the commercial vehicles resulting in imbalance of vehicle weight and uncontrolled driving situations. However, this fact is contradicted by a similar study by Chen and Chen (2011) where tire-defects decrease the likelihood of possible of non-incapacitating injuries. Single trailing unit decreases the major injury categories (i.e., fatality, incapacitating injuries, non-incapacitating injuries, possible injuries by 35.3%, 36.2%, 20.4%, and 39.8%, respectively).

Crash Characteristics or Mechanism: Departing the roadway (by left or right side of road way) increases 12.2% likelihood of fatalities which is also supported by a study by Chen and Chen (2011). Head-on collision also increases 11.5% likelihood of fatalities. This fact is supported by Chen and Chen (2011) through the variables driving on the wrong side or wrong way which might indicate its head-on impact with on-coming vehicles (e.g., wrong way driving). However, rollover situation increases 9.6% likelihood of incapacitating injuries which is complex in nature for multi-vehicle collisions. This fact is contradicted by Chen and Chen (2011) findings on truck overturn collisions. Rear-end collision increases 22.2% likelihood of incapacitating injuries and decreases the likelihood of non-injury collisions. Sideswipe in the same direction decreases 12.4% likelihood of non-incapacitating injuries. Number of vehicle involved in the multi-vehicle collisions increases 55.4% likelihood of possible injuries which is supported by a similar study by Chen and Chen (2011) with the fact that more than three vehicles involved in the collision increases both fatal and incapacitating and non-incapacitating or possible injuries. Speeding for the existing driving condition increases the likelihood of possible injuries which is supported by the fact that exceeding speed limit increases the likelihood of possible or non-incapacitating injuries in a study by Chen and Chen (2011). Lane changing behavior increases the likelihood of non-injury collisions (i.e., property-damage-only) which is practical for the situation of multiple vehicle interactions. As found in the random parameter ordered probit model, vehicles struck by other vehicles as consequences of vehicular interaction reduces the likelihood of severe injuries. Likewise, driving or going straight keeping the lane also reduces likelihood of severe injuries.

Table 6 Average Direct Pseudo-Elasticities of all variables in Multi-Vehicle Injury Model

Variables	Elasticity (%)				
	PDO/No Injury	Possible Injury	Non-incapacitating	Incapaci-tating	Fatal
Vehicle maneuver during pre-crash situation (1 if left or right side departure, 0 otherwise)	-0.05	-0.07	-0.05	-0.12	12.15
Light condition of street (1 if dark, 0 otherwise)	-0.23	-0.26	-0.13	-0.36	50.26
Orientation of vehicle at the time of crash (1 if head-on, 0 otherwise)	-0.05	-0.06	-0.03	-0.04	11.47
Time of the day (1 if 2 pm in the afternoon, 0 otherwise)	-0.14	-0.17	-0.08	-0.13	30.06
Driver's attention level at the time of impending crash (1 if distraction or in attention, 0 otherwise)	-0.16	-0.19	-0.10	7.75	-0.15
Time of the day (1 if 5 am in the morning, 0 otherwise)	-0.49	-0.54	-0.32	23.86	-0.83
Vehicular factors (1 if tire-related malfunction, 0 otherwise)	-0.11	-0.11	-0.08	5.24	-0.11
Orientation of vehicle at the time of crash (1 if rear-end, 0 otherwise)	-0.44	-0.66	-0.30	22.17	-0.46
The most harmful event in crash consequences (1 if rollover, 0 otherwise)	-0.20	-0.10	-0.19	9.60	-0.13
Orientation of vehicle at the time of crash (1 if sideswipe in the same direction, 0 otherwise)	0.47	0.41	-12.37	0.32	0.37
Occupants' use of available vehicle restraints (1 if no restraint used, 0 otherwise)	-0.42	-0.55	11.48	-0.63	-0.58
Time of the day (1 if 4 am in the morning, 0 otherwise)	-0.13	-0.12	3.49	-0.33	-0.18
Months of year (1 if summer months (June to August), 0 otherwise)	-0.41	-0.48	10.96	-0.49	-0.40
Gender of the occupants (1 if male, 0 otherwise)	2.33	-38.28	1.17	2.56	2.15
Drivers' working/residing place according to license record (1 if Texas, 0 otherwise)	-0.78	12.80	-0.41	-0.92	-0.67
Number of vehicles involved in the crash	-3.37	55.42	-1.72	-4.06	-3.05
Speed-related factor in crash (1 if speed as a factor, 0 otherwise)	-0.22	3.67	-0.13	-0.31	-0.24
Road surface condition (1 if wet, 0 otherwise)	-0.57	9.43	-0.29	-0.62	-0.55
Vehicle maneuver just prior to impending crash (1 if changing lane, 0 otherwise)	0.27	-2.57	-1.19	-1.99	-2.24
Light condition of street (1 if the surrounding area is dark but outside is lighted, 0 otherwise)	0.57	-4.86	-2.77	-5.88	-2.61
Driver's attention level at the time of impending crash (1 if sleepy, 0 otherwise)	-0.09	0.92	0.21	1.04	0.81
Trailing unit when the crash occurred (1 if one trailer, 0 otherwise)	4.39	-39.80	-20.37	-36.19	-35.30
Orientation of vehicle at the time of crash	-1.08	10.05	4.64	9.22	6.71

Variables	Elasticity (%)				
	PDO/No Injury	Possible Injury	Non-incapacitating	Incapaci-tating	Fatal
(1 if rear-end, 0 otherwise)					
Alignment of highway section (1 for curved section, 0 otherwise)	-0.31	2.99	1.14	2.41	2.44

CONCLUSION

Utilizing national wide representative GES crash database, we investigated two discrete outcome models – random parameter ordered probit because of ordinal characteristics of injury scale (as we followed KABCO scale) and mixed logit model because of methodological flexibility such as each injury outcome has individual utility functions and independent of irrelevant alternatives (IIA). The IIA property in mixed logit model provides the flexibility of variables within each of particular injury outcomes being independent as well as the between the outcomes (Jones and Hensher, 2007). Random parameter ordered probit provides the indication of more and less severe injury outcomes based on the sign of the variables which is not the case of mixed logit where more or less severe injury outcomes are distinctly declared as individual injury outcome in the utility function set up. The results of both models are presented here (Table 3 and Table 4) as well as their comparative statistical performance. The parameter estimates and statistical comparison clearly indicates that mixed logit model is superior to random parameter ordered probit model.

Few crucial human factors were indentified in this study which is worth mentioning such as distracted and sleepy driving, being female occupants (i.e., drivers or passengers), drivers residing or working in the state of Texas, not using seat-belt, as listed in the Empirical Results section clearly indicate potential dangers of being seriously injured in multi-vehicle large truck collisions. Turning to next important factors – road and environmental factors such as dark condition, time of day such as 2 pm in the afternoon, 4 to 5 am in the morning increase severe injuries. On the other hand, curved segments and wet surface condition increases minor injuries. Turning to vehicular factors such as tire related defects increases severe injury whereas, single trailing unit increases non-injury (i.e., PDO). Last factors which is part of human factors as the maneuvers are mainly executed by the drivers at the impending or pre-crash situations such as departing roadway, rear-end collision, head-on collision, rollover increases the severity of injuries. On the other hand, sideswipe (same direction), number of vehicles involved in the collisions, speeding for the condition such as unfavorable weather or existing traffic condition increases minor injuries in multi-vehicle large truck collisions.

Although we used same dataset and attempted with similar variables in both models, mixed logit model captured more variables than random parameter ordered probit model. These variables explained the multi-vehicle large trucks collisions. These variables could be categorized into factors: human factor – sleepy driving; road and environment factors – 2 pm, 4 to 5 am, wet surface, curved segments, dark but lighted condition; vehicular factor – tired related defects; and crash mechanism – head-on and rear-end collision, number of vehicles involved in the collisions.

Although GES dataset does not contain any traffic information (such average annual daily traffic or vehicle-mile travelled, etc.), proxy variables such as time of day (2 pm, 4 to 5 am of the day), month of year (June to August) were considered in the mixed logit and random parameter ordered probit model to capture traffic condition at the time of crash. In addition, this research also leads to further research into single vehicle collisions using the same database. It is worth investigating the contributing factors for single and multi-vehicle collisions involving large trucks and understands their role and differences in leading crashes on US highways.

REFERENCES

Abdel-Aty, M. "Analysis of Driver Injury Severity Levels at Multiple Locations using Ordered Probit Models," *Journal of Safety Research* 34 (2003): 597–603.

Anastasopoulos, P.Ch., and Mannering, F.L. "A Note on Modeling Vehicle-Accident Frequencies with Random Parameter Count Models," *Accident Analysis and Prevention*, 41(1) (2009):153-159.

Bhat, C. "Quasi–random Maximum Simulated Likelihood Estimation of the Mixed Multinomial Logit Model," *Transportation Research Part B*, 17(1) (2001): 677-693.

Bhat, C. Simulation Estimation of Mixed Discrete Choice Models using Randomized and Scrambled Halton Sequences. *Transportation Research Part B*, 37(1) (2003): 837-855.

Brownstone, D., and Train, K., "Forecasting New Product Penetration with Flexible Substitution Patterns," *Econometrica*, 89 (1999): 109-129.

Chen, F., and Chen, S., "Injury severities of truck drivers in single- and multi-vehicle accidents on rural highways," *Accident Analysis and Prevention*, 43(5) (2011): 1677-1688.

Chistoforou, Z., Cohen, S., Karlaftis, M. G., "Vehicle Occupant Injury Severity on Highways: An Empirical Investigation," *Accident Analysis and Prevention*, 42(6) (2010): 1606 –1620.

Federal Highway Administration. 2010. Highway statistics, 2008. Washington, DC: US Department of Transportation.

Fan, Y., and Lord, D. Investigating the Effects of Underreporting of Crash Data on Three Commonly Used Traffic Crash Severity Models: Multinomial Logit, Ordered Probit and Mixed Logit Models, *90th Transportation Research Board Annual Meeting*, (2011): TRB DVD (Paper# 11:1930).

Geweke, J., Keane, M., Runkle, D. Alternative computational approaches to inference in the multinomial probit model. *The Review of Economics and Statistics* 76(4), (1994): 609-632.

Greene, W., 2007. Limdep, Version 9.0. Econometric Software, Inc., Plainview, New York.

Gray, R.C., Quddus, M.A., Evans, A. "Injury Severity Analysis of Accidents involving Young Male Drivers in Great Britain," *Journal of Safety Research*, 39 (2008): 483–495.

Halton, J. "On the Efficiency of Certain Quasi-Random Sequences of Points in Evaluating Multi-Dimensional Integrals," *Numerische Mathematik*, 2 (1960): 84–90.

Insurance Institute for Highway Safety, Highway Safety Research and Communications. http://www.iihs.org/research/fatality_facts_2009/largetrucks.html.

Islam, M.B. and Hernandez, S. Large Truck involved Crashes: An Exploratory Injury Severity Analysis, 91st Transportation Research Board (CD-ROM), Washington D.C., USA (2012): (Paper#12-1961).

Jones, S., Hensher, D. "Evaluating the Behavioral Performance of Alternative Logit Models: An Application to corporate Takeovers Research," *Journal of Business Finance Accounting*, 34 (2007): 1193–1220.

Kim, J.-K., Ulfarsson, G., Shankar, V., Mannering, F. "A note on modeling pedestrian injury severity in motor vehicle crashes with the mixed logit model," Accident Analysis. And Prevention, 42(6), (2010): 1751–1758.

Kockelman, K.M., Kweon, Y.-J., "Driver Injury Severity: An Application of Ordered Probit Models," *Accident Analysis and Prevention*, 34(3) (2002): 313–321.

Lee, C., Abdel-Aty, M. "Comprehensive Analysis of Vehicle–Pedestrian Crashes at Intersections in Florida," *Accident Analysis and Prevention*, 37 (2005): 775–786.

Lee, L. "On Efficiency of Method of Simulated Moments and Maximum Simulated Maximum Estimation of Discrete Response Models," *Econometric Theory*, 8(3) (1992): 518–552.

Limdep, NLOGIT Version 4.0, Econometric Software, Planeview, NY.

McKelvey W., and Zavoina, T. "A Statistical Model for Analysis of Ordinal Level Dependent Variables," Journal of Mathematical Sociology, 34 (1975): 103–120.

McFadden, D., and Train, K. "Mixed MNL Models for Discrete Response," *Journal of Applied Econometrics*, 15 (2000): 447-470.

McFadden, D. Econometric models of probabilistic choice. In: Manski, D., McFadden (Eds.), A Structural Analysis of Discrete Data with Econometric Applications. The MIT Press, Cambridge, MA, 1981.

Milton, J. C., Shanker, V. N., Mannering, F. "Highway Accident Severities and the Mixed Logit Model: An Exploratory Empirical Analysis," *Accident Analysis and Prevention*, 40(1) (2008): 260-266.

National Automotive Sampling System (NASS) General Estimates System (GES), Analytical User's Manual 1988-2005. http://www-nrd.nhtsa.dot.gov/Pubs/AUM05.pdf.

NHTSA, 2008. Traffic Safety Facts 2007: Large Trucks, DOT HS 810 805. http://www-nrd.nhtsa.dot.gov/Pubs/810989.PDF.

O'Donnell, C.J., Connor, D.H., "Predicting the Severity using Models of Ordered Multiple Choice," *Accident Analysis and Prevention*, 28(6) (1996): 739–753.

Pai, C.-W., Saleh, W. "Modelling Motorcyclist Injury Severity by Various Crash Types at T-junctions in the UK," *Safety Science*, 46 (2008): 1234–1247.

Quddus, M.A., Noland, R.B., Chor Chin, H. "An Analysis of Motorcycle Injury and Vehicle Damage using Ordered Probit Models," *Journal of Safety Research*, 33(2002): 445–462.

Revelt, D., and Train, K., "Mixed Logit with Repeated Choices: Households' Choice of Appliance Efficiency Level," *Review of Economics and Statistics*, 80(4) (1997): 647-657.

Stern, S. "Simulation-based Estimation," *Journal of Economic Literature*, 35(4) (1997): 2006-2039.

Statistical Analysis System, SAS® 9.2 for Windows. SAS Institute, Inc., Cary, NC.

Train, K. Mixed Logit Models for Recreation Demand. In: Kling, C., Herriges, J. (Eds.), Valuing the Environment Using Recreation Demand Models. Elgar Press, NY, 1997.

Train, K. Discrete Choice Methods with Simulation. Cambridge University Press, Cambridge, UK, 2003.

Walker, J., Ben-Akiva, M., "Generalized Random Utility Model," *Mathematical Social Sciences*, 43(3) (2002): 303–343.

Washington, S.P., Karlaftis, M.G., and Mannering, F.L. Statistical and Econometric methods for Transportation Data Analysis, 2nd Ed. Chapman & Hall/CRC, 2010.

Xie, Y., Zhang, Y., Liang, F. "Crash Injury Severity Analysis Using Bayesian Ordered Probit Models," *Journal of Transportation Engineering*, 135(2009): 18-25.

Zaloshnja, E., and Miller, T.R. "Costs of Large Truck-involved Crashes in the United States," *Accident Analysis and Prevention*, 36(5) (2004): 801-808.

Zaloshnja, E., and Miller, T.R., 2006. Unit Costs of Medium and Heavy Truck Crashes, Final Report for Federal Motor Carrier Safety Administration and Federal Highway Administration, 400 Seventh Street, SW, Washington, DC.

Zhu, X., Srinivasan, S. "A Comprehensive Analysis of Factors Influencing the Injury Severity of Large-Truck Crashes," *Accident Analysis and Prevention*, 43(1)(2011): 49-57.

Zajac, S., Ivan, J.N. "Factors Influencing Injury Severity of Motor Vehicle-Crossing Pedestrian Crashes in Rural Connecticut," *Accident Analysis and Prevention*, 35(2002): 369–379.