Analysis of Injury Severity of Large Truck Crashes in Work Zones

Mohamed Osman [a], Rajesh Paleti [b], Sabyasachee Mishra [a,c1], Mihalis M. Golias [a,c]

[a] Department of Civil Engineering, University of Memphis, 3815 Central Avenue, Memphis, TN 38152, United States

[b] Department of Civil & Environmental Engineering, Old Dominion University, 135 Kaufman Hall, Norfolk, VA 23529, United States

[c] Intermodal Freight Transportation Institute, University of Memphis, Memphis, TN 38152, United States

Abstract

Work zones are critical parts of the transportation infrastructure renewal process consisting of rehabilitation of roadways, maintenance, and utility work. Given the specific nature of a work zone (complex arrangements of traffic control devices and signs, narrow lanes, duration) a number of crashes occur with varying severities involving different vehicle sizes. In this paper we attempt to investigate the causal factors contributing to injury severity of large truck crashes in work zones. Considering the discrete nature of injury severity categories, a number of comparable econometric models were developed including multinomial logit (MNL), nested logit (NL), ordered logit (ORL), and generalized ordered logit (GORL) models. The MNL and NL models belong to the class of unordered discrete choice models and do not recognize the intrinsic ordinal nature of the injury severity data. The ORL and GORL models, on the other hand, belong to the ordered response framework that was specifically developed for handling ordinal dependent variables. Past literature did not find conclusive evidence in support of either framework. This study compared these alternate modeling frameworks for analyzing injury severity of crashes involving large trucks in work zones. The model estimation was undertaken by compiling a database of crashes that (1) involved large trucks and (2) occurred in work zones in the past 10 years in Minnesota. Empirical

[1] Corresponding author.: Tel.: +1 901 678 5043
E-mail addresses: mosman@memphis.edu (M. Osman), rpaleti@odu.edu (R. Paleti), smishra3@memphis.edu (S. Mishra), mgkolias@memphis.edu (M.M. Golias)

findings indicate that the GORL model provided superior data fit as compared to all the other models. Also, elasticity analysis was undertaken to quantify the magnitude of impact of different factors on work zone safety and the results of this analysis suggest the factors that increase the risk propensity of sustaining severe crashes in a work zone include crashes in the daytime, no control of access, higher speed limits, and crashes occurring on rural principal arterials.

Keywords: large truck, work zones, injury severity, multinomial logit, ordered logit, generalized ordered logit

1. Introduction

Work zone safety is a major concern for the Federal Highway Administration (FHWA), State Departments of Transportation (DOTs), and the public. Over the last 30 years, the total lane miles in the US have increased by 7.4% whereas the Vehicle Miles Travelled (VMT) increased by 86% (FHWA, 2012). With increased VMT, work zone fatalities and injuries have also increased. Nationally, there were 87,606 work zone crashes in 2010 which is approximately 1.6% of the total number of roadway crashes. More than 20,000 workers were injured in work zones in 2010. In the same year, work zone crashes resulted in 37,476 injuries which equates to approximately four injuries every hour. In 2010, there were 514 fatal crashes resulting in 576 fatalities in work zones, which equates to approximately one fatality every 15 hours (FHWA, 2010). Work zones have unique traffic conditions that are different from other crash locations and thus warrant studies that focus exclusively on these locations instead of pooling them with other locations.

Another key segment of crashes, that is of major concern both to the transportation officials and the trucking industry, are those involving large trucks. In 2012 alone, there were 317,000 large truck crashes in the US that resulted in 3,464 fatalities and 73,000 injuries (FHWA, 2014). In the same year, large trucks accounted for 8% of all vehicles involved in fatal crashes and 3% of vehicles involved in injury and property-damage-only (PDO) crashes (U.S. Department of Transportation, 2014). Although these percentages may not seem alarming at first glance, the economic impact could be substantial because large truck crashes incur high costs including high value goods, and higher travel delays associated with longer traffic incident durations. Moreover, the determinants of severity of crashes involving large trucks can be considerably different from crashes involving passenger cars and/or relatively smaller commercial fleet. So, it is important to focus exclusively on large truck crashes to understand the relative effect of different factors on their road safety.

The current study aims to contribute to the literature on work zone safety by exploring the characteristics of large truck crashes in work zones using a disaggregate-level analytical approach that focusses on each individual crash and associated set of potentially contributing factors. Specifically, the study examines the factors that impact the severity level of the most severely injured individual involved in the crash, which essentially marks the overall severity level of the crash. Understanding large truck crash severity characteristics in work zones will be a steppingstone in enabling practitioners, designers, and DOT officials to mitigate the severity of such type of crash. The findings of this study have important implications in the work zone safety field, education of motorists, training of truck drivers, and traffic regulation and control. Designers of roadway work zones will be able to implement effective safety measures that will allow DOT officials to better manage the safety of a work zone through learning about the important factors influencing crashes involving large trucks.

The remainder of this paper is structured as follows. A literature review is presented in the next section followed by the econometric framework describing the methodology of the different models developed in this paper. The data section discusses the dataset utilized and the final estimation sample assembly process. The empirical analysis section presents a detailed overview of the estimation results, statistical measures of fit, and elasticity effects. Finally, the conclusion section provides an overall summary of this research along with major findings and future scope of research.

2. Literature Review

Several research studies have been conducted to analyze the severity of crashes involving large trucks (Chang and Mannering, 1999; Duncan et al., 1998; Islam and Hernandez, 2013; Li and Bai, 2009; Pahukula et al., 2015; Qi et al., 2013; Wang and Shi, 2013; Wang et al., 2010). The overview of the literature indicates that there is a vast body of research examining the factors affecting the severity of large truck-involved accidents on both crash-level and occupant-level. The literature presented in this paper is primarily focused on injury severity of large trucks in work zones at the crash-level to obtain insights and to help to meet the goal of this research. However, occupant-level injury severity studies are imperative in the context of work zone safety and comprehensively presented in the literature (Chang and Chien, 2013; Chen and Chen, 2011; Dong et al., 2015; Khorashadi et al., 2005; Lemp et al., 2011; Mooradian et al., 2013; Wong et al., 2011; Zhu and Srinivasan, 2011a, 2011b) and will not be reviewed herein.

The past literature can be grouped under three categories – (1) those that focus exclusively on large truck crash severity modeling, (2) those that focus on injury severity in the context of work zone safety, and (3) those that focus both on large truck crash severity and work zone safety combined. In this section we present a review of the crash-level literature that specifically pertained to injury severity of crashes involving large trucks, work zones, or both. The econometric framework comparisons utilized in this study have been recently used by other researchers in the context of injury severity analysis to evaluate alternate discrete outcome frameworks for modeling crash injury severity (Yasmin and Eluru, 2013). Sample size requirements were evaluated by comparing three commonly crash severity models (Ye and Lord, 2014). Another study has evaluated alternate discrete choice frameworks for modeling ordinal discrete variables but not necessarily in the context of injury severity (Eluru, 2013). A discrete choice model comparison was applied to investigate cyclist injury severity in automobile-involved bicycle crashes (Chen and Shen, 2016). Pedestrian Injury Severity in New York City was also examined using alternative ordered response frameworks (Yasmin et al., 2014). To our knowledge, this is the first application of such a comprehensive set of discrete choice models in the context of work zone safety. A brief overview of past literature in these three categories follows in the next three subsections.

2.1. Large truck crash severity

A variety of discrete choice models were used in the literature to analyze large truck crash severity. For example, assessing the severity of truck crashes on a freeway network using a hierarchical regression model indicated that the presence of ramp, freeway segment length, and weather conditions were important factors affecting truck safety performance (Wang and Shi, 2013). Utilizing nested logit models to investigate the severity in truck and non-truck crashes, risk factors that are unique to large trucks were identified. Variables that increased injury severity for large trucks were higher speed limits, vehicles making right or left turns, and rear-end collisions (Chang and Mannering, 1999). Using a random-parameter ordered probit model allowed the identification of the differences between random and fixed factors affecting the severity outcome. It was found that the severity level is highly influenced by complex interactions between factors, and that the effects of some variables can vary across observations (Islam and Hernandez, 2013). Investigating rear-end large truck crashes using an ordered probit model indicated that darkness, high speed differential between vehicles and trucks, higher speed limits, wet surfaces on a grade, a car struck to the rear, and alcohol increased crash severity while snow and ice, congested roads, and station wagon decreased the likelihood of a severe crash (Duncan et al., 1998). An exploratory study utilized a mixed logit model to analyze injury severity of crashes involving large trucks on Texas highways which revealed that time-of-day (12-6 AM), summer time (June-August), clear weather, rural areas, and 4-lane roadways were all contributing factors to higher likelihood of higher injury severity levels (M. Islam and Hernandez, 2013). Another study also used mixed logit models to estimate the effect of time of day on injury severity of large truck crashes in urban areas (Pahukula et al., 2015). The study uncovered major differences both in the combination of variables and their magnitude of impact on the severity outcomes across different time periods. Among different explanatory variables used in the study, the effects of traffic flow, lighting road surface conditions, time of year, and percentage of trucks were found to vary by time period (Pahukula et al., 2015). In recent years, mixed logit models have generally gained attention within the discrete choice modeling literature due to their flexibility in allowing variations over data observations as compared the restrictions imposed by standard logit models. This modeling technique has been utilized in previous large truck literature, but not necessarily within the context of injury severity (Romo et al., 2014).

2.2. Work zone crash severity

A work zone crash is defined as a crash that occurred in an area comprising a work zone as per defined by the Manual of Uniform Traffic Control Devices (MUTCD). Specifically, for the purpose of this study, a work zone extends from the “advanced warning area” until the “termination area”. There is some literature that focused specifically on crashes in work zones. For instance, one study used the ordered probit model to analyze severity of rear-end crashes in work zones. The study found that alcohol, night hours, pedestrians, roadway defects, truck-involvement, and the number of vehicles involved increased crash severity, while careless backing, stalled vehicles, slippery surfaces, and misunderstanding flagging signals resulted in less severe injuries in the event of a crash (Qi et al., 2013). However, there is no consensus on these findings in the safety literature. Other studies that used similar discrete choice modeling methods found slightly contradicting results (Wang et al., 2010). Another study by Wong et al., 2011) examined factors influencing injury severity of highway workers in work zone intrusion crashes using multiple correspondence analysis, Cox proportional hazard regression, logistic regression, and Poisson regression models and found that work zone location and duration, time of the day, and type of activity performed by workers were the most significant factors impacting severity outcomes.

2.3. Large truck crash severity in work zones

Most of the crash severity literature to date provide only basic information in terms of the large truck-involvement in a work zone crash (Li and Bai, 2008; Qi et al., 2013; Wang et al., 2010). Such studies used large truck-involvement in a work zone crash as a binary explanatory variable in severity models. There is only one study in the literature that modeled injury severity of crashes involving large trucks in work zones. Khattak and Targa, 2004 have modeled injury severity and total harm in work zone crashes involving large trucks by assigning an economic cost for the different severity levels. The study found that, on average, large truck crashes that occurred on two-way undivided roads, roads with higher speed limits, and in the proximity of work zones tend to be more severe than other crashes. Given the relatively sparse literature on work zone crashes involving large trucks, the current study aims to develop improved tools that can provide better insights by using new econometric methods that were developed recently. Specifically, the current study compared the performance of alternate modeling frameworks in identifying significant factors affecting the severity of large truck crashes in work zones.

3. Econometric Framework

The modeling methods typically used to analyze crash data pertaining injury severity can be grouped into two categories – unordered and ordered. In the unordered modeling framework, the observed severity outcome is assumed to be the outcome with the highest latent severity function value (there is one severity function corresponding to each severity outcome). Each of the latent severity functions is specified as a linear function of different crash factors with a stochastic component to account for all unobserved factors that influence the corresponding severity outcome. The coefficients in all the severity functions constitute the set of parameters that are estimated using inference methods such as the maximum likelihood (ML) estimation approach. In the ordered framework, on the other hand, a single latent propensity function is assumed to be translated into the observed severity outcome depending on the value of the propensity function relative to threshold parameters (number of thresholds = number of possible severity outcomes – 1). The latent propensity function is specified as a function of different factors along with a stochastic component to account for all unobserved factors that influence crash severity. The parameters in the single propensity equation and the thresholds constitute the set of parameters that are estimated using methods such as the ML. Earlier comparison studies for analyzing ordinal discrete outcomes (not necessarily in the context of severity analysis) found that the unordered framework fits data better than ordinal models because of the flexibility provided by additional parameters in the unordered models. However, a study by Eluru *et al*., (2008) developed generalized ordered models that allow parameterization of the threshold parameters providing additional flexibility to the ordinal models (Eluru et al., 2008). So, it is not surprising that a recent comparison analysis of unordered and ordered frameworks that considers generalized version of ordered models found minor differences between the two models (Anowar et al., 2014). So, it is imperative that researchers compare and choose the best method specific to the empirical context of interest. This section describes the two modeling frameworks and their generalized variants used in this study.

3.1. Unordered modeling framework

Let *i* be the index for the injury severity outcome (1 = "no injury", 2 = "injury", and 3 = "serious injury") and *n* be the index for crash. Also, let *I* denote the total number of severity outcomes (which is 3 in the current empirical context) and *N* denote the total number of crashes in the dataset. In this study, a linear-in-parameter specification was adopted for the deterministic part of U_{in} as follows: $U_{in} = \boldsymbol{\beta}_i'\boldsymbol{X}_{in} + \varepsilon_{in}$ where $\boldsymbol{X}_{in}$ is a $K_i \times 1$ vector of exogenous covariates (including crash factors, work zone attributes, environmental, and roadway conditions), $\boldsymbol{\beta}_i$ is the corresponding $K_i \times 1$ vector of coefficients and ε_{in} denotes all the unobserved factors that influence the severity function for outcome *i* in crash *n*. As discussed earlier, in the unordered framework, the observed severity outcome is the severity outcome with the highest latent severity function value. So, the probability that crash *n* sustains severity outcome *i*, $P_n(i)$ is given by:

$$P_n(i) = P\left(\boldsymbol{\beta}_i'\boldsymbol{X}_{in} + \varepsilon_{in} \geq \boldsymbol{\beta}_j'\boldsymbol{X}_{jn} + \varepsilon_{jn}\right) \forall\, j \neq i$$

3.1.1. Multinomial logit (MNL) model

In the MNL model, the stochastic components ε_{in} in the latent severity functions U_{in} are assumed to be independent and identically distributed (*i.i.d.*) across different severity outcomes and crashes. Moreover, the identical distribution is assumed to be standard type-1 extreme value distribution (also referred to as Gumbel distribution). Given these assumptions on the stochastic term ε_{in}, $P_n(i)$ can be derived to be:

$$P_n(i) = \frac{\exp(\boldsymbol{\beta}_i'\boldsymbol{X}_{in})}{\sum_{\forall I} \exp\left(\boldsymbol{\beta}_j'\boldsymbol{X}_{jn}\right)}$$

The $\sum_{i=1}^{I} K_i$ parameters in the MNL model were estimated by maximizing the log-likelihood function obtained by taking the natural logarithm of the product of probabilities of observed severity outcomes given by Equation (2) as follows:

$$LL = \sum_{n=1}^{N} \left(\sum_{i=1}^{I} \delta_{in} \right)$$

where δ_{in} is defined as 1 if the observed severity outcome for crash n is i and zero otherwise.

3.1.2. Nested logit (NL) model

The MNL model has the Independence of Irrelevant Alternatives (IIA) property which implies that changes in conditions that influence one severity outcome do not change the relative probabilities of other severity outcomes. This can be a strong restrictive assumption in the current empirical context given that severity data is ordinal in nature with potentially strong correlations between successive severity outcomes. Past literature found evidence for correlation among unobserved effects to be present (Shankar et al., 1996), while other research has not (Shankar and Mannering, 1996). Assuming the IIA property to hold in cases when it is violated can produce incorrect parameter estimates because of specification errors. The NL model that relaxes the IIA assumption by allowing correlation in unobserved factors of subsets of alternatives is more suited for such scenarios (Shankar et al., 1996). In this study, alternate two-level nesting structures that group all the severity outcomes into S mutually exclusive and exhaustive nests B_s each with nesting parameter ρ_s $(0 < \rho_s \leq 1)$ were estimated. The probability of severity outcome i that belongs to nest B_r can be obtained as the product of conditional probability of the outcome i within the nest B_r and the probability of the nest B_r among all possible nests B_s $s \in \{1,2, \ldots .S\}$ as follows:

$$P_n(i) = \frac{e^{\frac{\beta_i' X_{in}}{\rho_r}}}{\sum_{k \in B_r} e^{\frac{\beta_k' X_{kn}}{\rho_r}}} \times \frac{e^{\rho_r IV_r}}{\sum_{s=1}^{S} e^{\rho_s IV_s}} \text{ where Inclusive Value } (IV_s) = LN\left[\sum_{k \in B_s} e^{\frac{\beta_k' X_{kn}}{\rho_s}}\right]$$

3.2. Ordered response framework

3.2.1. Ordered logit (ORL) model

As discussed earlier, in the ordinal framework, latent propensity y_n^* is translated into observed severity outcomes by threshold parameters. This study adopted a linear-in-parameter specification for the observed part of y_n^* and a standard logistic distribution that is *i.i.d.* across crashes for the stochastic component ε_n . The equation system for the ORL model is (McKelvey and Zavoina, 1975):

$$y_n^* = \boldsymbol{\beta}'\boldsymbol{X}_n + \varepsilon_n$$

$$\begin{aligned} P_n(i) &= P(\psi_{i-1} < y_n^* < \psi_i) \\ &= P(\psi_{i-1} < \boldsymbol{\beta}'\boldsymbol{X}_n + \varepsilon_n < \psi_i) \\ &= P(\psi_{i-1} - \boldsymbol{\beta}'\boldsymbol{X}_n < \varepsilon_n < \psi_i - \boldsymbol{\beta}'\boldsymbol{X}_n) \\ &= F(\psi_i - \boldsymbol{\beta}'\boldsymbol{X}_n) - F(\psi_{i-1} - \boldsymbol{\beta}'\boldsymbol{X}_n) \end{aligned}$$

where $\boldsymbol{X}_n$ is $K \times 1$ vector of covariates and $\boldsymbol{\beta}$ is the corresponding $K \times 1$ vector of coefficients; $\psi_i's$ are threshold parameters; $\psi_0 = -\infty \; and \; \psi_{I+1} = \infty$; $F(.)$ is the standard logistic cumulative distribution function. The model structure requires that the thresholds to be strictly ordered for the partitioning of the latent risk propensity measure into the ordered injury severity categories$(i.e., -\infty < \psi_1 < \psi_2 < \cdots < \psi_{I-1} < \infty)$. The parameters in the ORL model ($\boldsymbol{\beta}$ and $\psi_i's$) were estimated using the ML inference method.

3.2.2. Generalized ordered logit (GORL) model

One of the restrictive assumptions of the standard ORL model is that it assumes that the threshold parameters do not vary across different crashes. Eluru *et al.*(2008) relaxed this assumption by parameterizing the thresholds as a function of exogenous factors providing additional flexibility to the model (Eluru et al., 2008). The structure of the GORL follows the same structure of the ORL in Equation (5) except for ψ parameters which are now subscripted by index n to reflect that these parameters will vary across crashes (Eluru et al., 2008; Romo et al., 2014).

$$y_n^* = \boldsymbol{\beta}'\boldsymbol{X}_n + \varepsilon_n$$

$$\begin{aligned} P_n(i) &= P\left(\psi_{n,i-1} < y_n^* < \psi_{n,i}\right) \\ &= F\left(\psi_{n,i} - \boldsymbol{\beta}'\boldsymbol{X}_n\right) - F\left(\psi_{n,i-1} - \boldsymbol{\beta}'\boldsymbol{X}_n\right) \end{aligned}$$

To ensure strict ordering of thresholds, the following parameterization was adopted:

$$\psi_{n,i} = \psi_{n,i-1} + exp(\alpha_i + \boldsymbol{\gamma}_i' \boldsymbol{Z}_{ni})$$

where z_{ni} is a set of exogenous variables associated with the i^{th} threshold excluding the constant; $\boldsymbol{\gamma}_i$ is the corresponding vector of coefficients, and α_i is a parameter associated with injury severity level $i = 1,2,\ldots,I-1$. $\psi_{n,1}$ is specified as $exp(\alpha_1)$ for identification reasons. The ORL model can be obtained from the GORL model by imposing the constraints that $\gamma_i = 0$ for all i.

4. Data

A dataset consisting of work zone crashes over 10 years (2003-2012) in Minnesota (MN) was collected from the Highway Safety Information System (HSIS). Two main datasets were obtained and merged. The first was the “accident file”, containing variables such as crash time, location, roadway condition, crash type, traffic control, and weather conditions. The second was the “road file”, containing basic characteristics of the roadway segment where the crash occurred such as lane, shoulder and median widths, speed limit, and several geometric design variables.

For the purposes of this study, only crashes involving at least one large truck were considered as truck-related crashes. The dataset contained 18,889 crashes in work zones with 15% involving large trucks (*i.e.*, 2,881 records were available for the analysis in this study). The crash severity level followed the KABCO injury severity scale where K=killed, A=incapacitating injury, B=non-incapacitating injury, C=possible injury, and O=no injury. The distribution of crashes by injury severity is presented in Tables 1.a and 1.b. Table 1.a shows the percentage of each severity category of the original data. Due to the low frequency of some of the severity levels, some of the severity categories were combined. The combined injury severity categories are shown in Table 1.b. Fatal, incapacitating, and non-incapacitating severity levels were combined into one severity level called “severe injury”. “Possible injury” which is referred to as “injury” and “no injury” categories were kept as is.

TABLE 1.a Frequency Distribution of Dependent Variable

Injury Severity Category	Count	(%)
Fatal (K)	19	0.66%
Incapacitating Injury (A)	29	1.01%
Non-Incapacitating Injury (B)	152	5.28%
Possible Injury (C)	435	15.10%
Property Damage (O)	2,246	77.96%
Total	2,881	100.00%

TABLE 1.b Frequency Distribution of Final Dependent Variable

Combined Injury Severity Category	Count	(%)
Serious Injury (K,A,B)	200	6.94%
Injury (C)	435	15.10%
No Injury (O)	2,246	77.96%
Total	2,881	100.00%

5. Empirical Analysis

Several categories of independent variables were considered in the empirical analysis to account for roadway, traffic, environmental, temporal, work zone, and crash characteristics. Table 2 indicates the frequency distribution of the explanatory variables. Roadway characteristics included functional class and geometric design factors. Functional class of each roadway was classified into one of the following types - "rural principal arterial", "urban principal arterial", "urban minor arterial", and "collectors, local systems or rural minor arterial". Geometric design factors included whether the road was curved or straight, number of lanes, and whether the roadway was curbed and access-controlled. Traffic characteristics included "speed limit" upstream of a work zone area. The effect of speed was captured using three categorical variables indicating whether speed limit was less than 35 mph, between 35 and 40 mph, between 45 and 50 mph, between 55 and 60 mph, or greater than 60 mph. Work zone immediate upstream speed limits were utilized in this research for each crash location. Environmental factors included wet surface and adverse weather (rain, fog, and snow). The impact of time of day was captured using three broad time categories - day (6 am - 6 pm), evening (6 pm - 12 am), and late night (12 am - 6 am). In addition to the time-of-day variables, an indicator variable for peak hours that denoted whether the crash occurred between 7-10 am or 4-7 pm was used. Work zone characteristics included the type of work zone (lane closure, shoulder or median work, lane shift or crossover, and intermittent/moving work zones). The crash work zone location indicated whether the crash has occurred in the proximity of advanced signage, work activity, transition, or termination areas. In addition to the variables listed above, an indicator variable for whether workers were present at the work zone was also tested during model estimation. Several geometric design variables were purposely omitted, such as (lane width, median width, shoulder widths), due to the fact that those types of variables are, most of the time, altered in a work zone depending on the nature and type of roadway work it is. This level of detailed work zone-specific geometric layout data was not available to the authors. Lastly, crash characteristics included the number of vehicles involved in the crash, truck type, and whether the crash occurred at a signalized intersection or on a bridge.

The final specifications for the presented models were based on a logical process of removing the statistically insignificant variables and combining other variables when their effects were statistically insignificant. The model estimation process was, in large part, guided by findings of past research and intuitiveness of the parameters estimated. Since work zones are naturally different than regular roadway segments in terms of roadway geometry, traffic controls and operational characteristics, the injury severity results in the current analysis are considered to be distinctive for work zones due to the special characteristics of roadways in work zones versus non-work zone areas. The final sample in the current study was narrowed down to those accidents that only occurred in a work zone while involving at least one large truck.

TABLE 2 Frequency Distribution of Explanatory Variable

Explanatory Variable	**(%)**	**Explanatory Variable**	**(%)**
Roadway		Time of the day	
Functional class		Day (6:00 AM - 6:00 PM)	6.4
Rural principal arterial	13.0	Evening (6:00 PM - 12:00 AM)	81.1
Urban principal arterial	57.7	Late night (12:00 AM - 6:00 AM)	12.5
Urban minor arterial	15.3	Work zone	
Other	14.0	Workers present	
Geometric design		Yes	42.4
Alignment		No	57.6
Curved	15.3	Work zone type	
Straight	84.7	Lane closure	36.6
Number of lanes		Lane shift/crossover	19.1
Two-lane	19.9	Shoulder or median	20.3
Multi-lane	80.1	Intermittent/moving	7.6
Curb		Other	16.4
Yes	37.5	Work zone location	
No	62.5	Advanced signs	8.0
Access control		Transition	18.8
Full	52.3	Activity	53.0
Partial	6.6	Termination	2.6
None	41.1	Other	17.6
Traffic		Crash	
Speed limit (mph)		Number of vehicles	
< 35	21.2	Single-vehicle	12.0
35 - 40	6.8	Multi-vehicle	88.0
45 – 50	12.4	Truck type	
55 - 60	50.7	Bus	7.9
65 - 70	8.9	2 axle 1 unit	15.6
Environmental		3+ axle 1 unit	11.7
Roadway surface condition		1 unit with trailer	6.5
Wet	15.8	Tractor-semitrailer	48.5
Dry	84.2	Other	9.8
Weather condition		Location	
Adverse	8.7	Signalized intersection	
Clear	91.3	Yes	15.1
Temporal		No	84.9
Peak hours		On-bridge	
Peak	34.7	Yes	6.5
Off-peak	65.3	No	93.5

6. Estimation Results

Table 3 presents the estimation results of the MNL, ORL, and GORL models. To test the validity of the IIA assumption of the MNL model, two-level nested logit (NL) models with two possible nesting structures with three severity outcomes were estimated. Neither nesting structure was found to be statistically sound as both nesting parameters did not fall between 0 and 1 (Manski and McFadden, 1981). So, the NL model was excluded from further analysis. The results corresponding to the MNL model consists of two columns labelled “injury”, and “serious injury”, while “no injury” category was chosen as the base category. The ORL model has one column corresponding to the variables in the propensity specification and two threshold parameters. The results corresponding to the GORL model are presented in two columns; the first column corresponds to the variables in the latent risk propensity (not including a constant) and the second column corresponds to the variables in the second threshold specification between the “injury” and “serious injury” outcomes. The respective t-values of the estimated coefficients are shown in parentheses. Table 3 also presents the initial log-likelihood value, the log-likelihood value at convergence, the Bayesian information criterion value (BIC), the McFaddenR^2, and the total number of crashes n for the three models.

6.1. Roadway characteristics

Rural principal arterials increased the likelihood of “serious injury” relative to “no injury” outcomes according to the MNL. Similar results were obtained from the ORL and GORL models. However, other functional class categories were also found to be statistically significant in the ordered response framework. To be specific, the OR models indicate that, on average, rural principal, urban principal, and minor urban arterials have higher risk propensity relative to rural minor arterials, collectors, and local systems.

Curved roadways were found to be associated with lower likelihood of sustaining “injury” but higher likelihood of “serious injury” relative to “no injury” outcomes in the event of a crash. This non-monotonic effect of road curvature is interesting. In some cases, it seems that steep curves are dangerous

TABLE 3 MNL, ORL, GORL model results

	MNL (Base Category: No Injury)		ORL	GORL	
Variable	**Injury**	**Serious Injury**	**Latent Propensity**	**Latent Propensity**	**Threshold: injury | serious injury**
Constant	-1.511 (-7.88)	-2.566 (-10.10)			
Roadway					
Functional class (base = collector, local system, rural minor arterial)					
Rural principal arterial	-	0.566 (2.81)	0.651 (3.22)	0.645 (3.21)	-
Urban principal arterial	-	-	0.454 (2.33)	0.434 (2.25)	-
Urban minor arterial	-	-	0.242 (1.29)	0.232 (1.24)	-
Geometric design					
Alignment (base = straight)					
Curved	-0.196 (-1.25)	0.283 (1.41)	-	-	-
Number of lanes (base = multi-lane)					
Two-lane	-	-0.445 (-1.95)	-	-	-
Curbed (base = no curb)					
Curb	-	-0.374 (-1.91)	-0.100 (-1.00)	-	0.214 (2.05)
Access-control (base = full control, and partial control)					
No control	0.263 (1.87)	0.950 (4.61)	0.654 (4.65)	0.612 (4.35)	-0.246 (2.42)
Traffic					
Speed limit (mph) (base = speed limit 45 to 60 mph)					
< 35 mph	-0.990 (-5.36)	-0.602 (-2.32)	-0.725 (-4.24)	-0.755 (-4.63)	-
35 - 40 mph	-0.366 (-1.61)	-0.651 (-1.78)	-0.398 (-1.93)	-0.404 (-1.98)	-
65 - 70 mph	-	0.579 (2.53)	0.284 (1.77)	0.232 (1.45)	-0.391 (-2.32)

TABLE 3 Continued

	MNL (Base Category: No Injury)		ORL	GORL	
Variable	**Injury**	**Serious Injury**	**Latent Propensity**	**Latent Propensity**	**Threshold: injury | serious injury**
Environmental					
Roadway surface condition (base = dry)					
Wet	-	-0.784(-2.25)	-0.369 (-2.03)	-0.348 (-1.92)	
Weather condition (base = clear)					
Adverse (rain, snow, fog, etc.)	-	0.455 (1.11)	0.230 (1.02)	0.213 (1.00)	-
Temporal					
Peak hours (base = off-peak)					
Peak	-0.187 (-1.63)	-	-0.162 (-1.63)	-0.156 (-1.57)	-
Time of the day (base = late night 12:00 AM - 6:00 AM)					
Day (6:00 AM - 6:00 PM)	0.400 (1.48)	0.972 (3.25)	0.567 (3.18)	0.493 (2.75)	-0.425 (-2.24)
Evening (6:00 PM - 12:00 AM)	0.277 (1.58)	-0.270 (-1.17)	-	-	-
Work Zone					
Workers (base = not present)					
Present	-	0.413 (2.60)	-	-	-
Work zone type (base = ln shift/crossover, intermittent/moving work zone)					
Lane closure	-0.279 (-2.42)	-0.349 (-2.05)	-0.236 (-2.21)	-0.245 (-2.30)	-
Shoulder or median	-	-	0.143 (1.20)	0.131 (1.11)	-
Work zone location (base = advanced signs, activity, termination, other areas)					
Transition area	-0.238 (-1.66)	-0.615 (-2.70)	-0.373 (-2.94)	-0.375 (-2.97)	-
Crash					
Number of vehicles (base = multi-vehicle)					
Single-vehicle	-0.372 (-1.98)	-	-0.151 (-1.01)	-0.191 (-1.27)	-0.470 (-2.80)
Truck type (base = bus, 2 axels 1 unit, other)					
3+ axle 1 unit truck	-	0.384 (1.77)	-0.160 (-1.14)	-	-
1 unit Truck with trailer	-	0.468 (1.70)	0.354 (2.05)	0.335 (1.94)	-
Truck tractor semitrailer	-0.132 (-1.21)	-	-	-	-

TABLE 3 Continued

	MNL (Base Category: No Injury)		**ORL**	**GORL**	
Variable	**Injury**	**Serious Injury**	**Latent Propensity**	**Latent Propensity**	**Threshold: injury | serious injury**
Location					
Signalized intersection (base = no signal)	-	-0.172 (-1.06)	-	-	-
On-bridge (base = not on-bridge)	0.347 (1.78)	-	0.234 (1.34)	0.233 (1.33)	-
Threshold coefficients (ORL, GORL)					
No Injury | Possible Injury			0.4883 (3.78)	0.4660 (3.56)	
Possible Injury | Serious Injury			0.3134 (6.64)	0.4863 (6.61)	
Log-Likelihood at zero	-1915.10		-1,915.10	-1,915.10	
Log-Likelihood at convergence	-1836.62		-1,862.50	-1,847.37	
BIC	3,912.21		3,876.35	3,870.00	
McFadden R^2	0.0410		0.0275	0.0354	
Number of observations	2881		2881	2881	

and can lead to severe outcomes in the event of crash and in few other cases, increased driver awareness and cautious driving while maneuvering curved roads reduces chances of injury (Lemp et al., 2011). Crashes on two-lane roadways tend to be less severe compared to crashes on multi-lanes roads. This finding is contrary to other studies that found that work zone crashes on two-lane roads were more severe; however these earlier studies focused on all crashes in work zones and did not control for the presence of a large truck (Li and Bai, 2009; Wang et al., 2010). All three models showed that crashes in work zones of curbed roadways were less severe compared to crashes on non-curbed roadways. It is important to note that, unlike in the ORL model in which the variable was present in the propensity equation, this variable was found to influence injury severity through the threshold parameter between the "injury" and "serious injury" outcomes. Specifically, a positive coefficient for 'curbed' roadway in the GORL threshold specification suggests wider translation region or higher likelihood of "injury" outcome and lower likelihood of "serious injury" outcome in the event of a crash.

Lack of access-control increased the likelihood of "injury" and "serious injury" relative "no injury" outcomes according to the MNL model. The positive coefficient values for the ORL and the GORL latent propensities showed similar results. Non-access-controlled roadways are likely to have more conflict points. The negative coefficient value for non-access-controlled roadways in the threshold specification of the GORL indicated an increased likelihood of "serious injury" relative to "injury" outcomes.

6.2. Traffic characteristics

All three models suggest that, on average, lower speed limits have lower risk propensity relative to higher speed limits. To be specific, the negative coefficients of speed limits of 40 mph or less were found to be associated with lower likelihood of sustaining "injury" and "serious injury" relative to "no injury" outcomes according to the MNL framework. Similarly, both OR models had negative coefficients in their propensity equations indicating the lower risk towards higher severity outcomes as compared to the base case of 45 to 60 mph. It was not surprising that the involvement of a large truck in a work zone crash while traveling at higher speeds essentially proposed a deadly combination. Speed limits of 65 mph or higher, on the other hand, indicated the higher likelihood of higher severity outcomes relative to the base case category explained by the positive coefficients of all three models. This variable was found to influence injury between the "injury" and "serious injury" outcomes through the threshold parameter according to the GORL model. A negative coefficient in the GORL threshold specification suggests lower likelihood of "injury" and higher likelihood of "serious injury" outcomes. Such a behavior was presented in earlier work zone crash severity literature; however a large truck involvement was not a factor (Li and Bai, 2009; Wang et al., 2010).

6.3. Environmental characteristics

Crashes on "wet surface" were associated with lower likelihood of "serious injury" relative to "no injury" in the MNL model. Similar results were obtained from the OR models. The ORL and GORL models indicate that roadways with wet surface have lower risk propensity relative to dry surface roadways. It seems as if truck drivers are more cautious driving at lower speeds and maintaining safe headways when driving on wet surface; such behavior has been suggested by past research (Chen and Chen, 2011; Duncan et al., 1998; Lemp et al., 2011; Zhu and Srinivasan, 2011a, 2011b). Crashes during "adverse weather" conditions were associated with higher likelihood of sustaining "serious injury" relative to "no injury" according to the variable positive coefficient in the MNL model. The "adverse weather" variable was also found to be statistically significant with similar results obtained in the OR models. The ORL and GORL indicate that "adverse weather" has higher risk propensity relative to clear weather conditions indicated by the positive coefficients in their risk propensity functions. This result is consistent with earlier large truck crash severity literature; however these studies did not control for crashes specifically in work zones (Chang and Mannering, 1999; Chen and Chen, 2011; Dong et al., 2015; Wang and Shi, 2013). Adverse weather is likely to be associated with poor sight distance and visibility.

6.4. Temporal characteristics

Travelling during "peak-hours" was found to be associated with lower likelihood of "injury" relative to "no injury" according to the MNL. Similar results were obtained in the OR framework. To be specific, the negative coefficients of the ORL and GORL models indicate lower risk propensity for traveling during peak-hours relative to non-peak hours. This is not a surprising result as traveling during peak-hours is typically congested leading to lower speeds, therefore reducing forceful impacts; such a result is consistent with past literature (Chang and Chien, 2013; Chang and Mannering, 1999; Duncan et al., 1998; M. Islam and Hernandez, 2013; Pahukula et al., 2015). All three models showed that crashes during daytime were more severe compared to other times of the day. In the MNL, the magnitude of the positive coefficients indicated the higher likelihood of "serious injury" relative to "injury" outcomes. The negative coefficient in the GORL threshold specification essentially showed similar results. The "evening" indictor, in the MNL, was associated with higher likelihood of sustaining "injury" but lower likelihood of "serious injury" relative to "no injury" outcomes in the event of a crash. It seems that traveling at night can lead to an injury crash but not severe enough to cause serious injuries. Past studies have found similar results (Islam and Hernandez, 2013). Crashes during evening times are likely associated with lower visibility and higher speeds due to lower traffic volumes.

6.5. Work zone characteristics

The presence of worker in a work zone was associated with higher likelihood of "serious injury" relative to "no injury" outcomes according to the MNL model. The MNL positive coefficient value for the "serious injury" outcome essentially indicated that workers on foot have greater odds of sustaining higher severity levels. Closing a lane or more in a work zone was found to be associated with lower likelihood of sustaining higher severity levels according to all three models. While the GORL failed to explain the effects of "lane closure" between the "injury" and "serious injury", the magnitude of the coefficients of both outcomes in the MNL indicated the lower likelihood of "serious injury" relative to "injury". Closing a lane or more is likely associated with the reduction of speed due to the combined traffic volumes into the functional lanes in a work zone. Work on shoulders or medians led to higher severity levels in the event of a crash indicated by the positive coefficients in both of the risk propensity equations of the OR models. This higher risk is likely associated with travelling adjacent to fully functional lanes where large trucks tend to drive at higher speeds compared to partially or fully closed lanes. Crashes in the transition area of a work zone were less likely to be severe as indicated by the negative coefficients in all three models. Drivers in the transition area have already passed through various advanced-warning and speed limit signs; the areas of a work zone that generally require lane changes and lane shifts, therefore motorist are likely to be already at lower speeds in those areas.

6.6. Crash characteristics

Crashes involving "single-vehicle" were found to be less severe according to the MNL and the risk propensity functions of both OR models; such a behavior was also suggested by earlier research (Qi et al., 2013). Interestingly, this variable had opposite effects in the GORL threshold equation between "injury" and "serious injury". Such a behavior suggests that although less involved vehicles can lead to lower likelihood of severe crashes, yet if an injury in fact occurred, the likelihood of "serious injury" is higher. Truck drivers are probably driving at higher speeds especially when not crowded by other vehicles in a work zone; therefore a sudden maneuver to change lanes or avoid workers could explain the opposite effects of the variable towards the lower and higher severity outcomes.

Crashes involving one-unit large trucks with three or more axles were found to be associated with higher likelihood of "serious injury" relative to the "no injury" outcomes indicated by the positive MNL coefficient for this variable. Past research have found similar result (Chen and Chen, 2011; Lemp et al., 2011). The more axles on a one-unit truck generally indicate heavier gross weight leading to forcible impacts. With a lower t-value in the ORL propensity equation, the negative coefficient value indicated the lower odds of higher severity levels; this behavior was also found in the literature (Chang and Chien, 2013; Dong et al., 2015; Khorashadi et al., 2005; Zhu and Srinivasan, 2011a). Given the occurrence of a crash; the MNL relative to the ORL models suggest that if an individual in fact has sustained an injury, it is severe.

Crashes involving one-unit trucks with trailers were more severe compared to other types of large trucks as indicated by the positive coefficients in all three models. Specifically, the MNL indicated the higher likelihood of "serious injury" relative to "no injury" outcomes as compared to buses, two-axle one-unit, and "other" types of large trucks. A trailer holds heavier cargo leading to higher severity levels. This result is consistent with the earlier research (Lemp et al., 2011; Zhu and Srinivasan, 2011a). The indicator of truck-tractor with a semitrailer was found to be associated with lower likelihood of "injury" relative to "no injury" according to the MNL negative coefficient; however, this result was associated with a lower statistical significance level. Previous large truck severity research suggested similar results; however these studies did not control for crashes specifically in work zones (Chen and Chen, 2011; Dong et al., 2015).

Signalized intersections were found to be associated with lower likelihood of "serious injury" relative to "no injury" outcomes as compared to non-signalized intersections according to the MNL model. Such a behavior was suggested by other studies; however these studies did not control for crashes specifically in work zones (Pahukula et al., 2015; Zhu and Srinivasan, 2011b). All three models had positive coefficients for the "on-bridge" variable which essentially showed that crashes occurring on a bridge in a work zone were more severe compared to crashes on non-bridged roadways. Generally, bridges are poor locations for a large truck to maneuver, especially in a work zone where lane, shoulder, and median widths are usually kept at a minimum. This results is consistent with past work zone crash severity literature; however large truck involvement was not controlled for in those studies (Qi et al., 2013).

6.7. In-depth analysis of large truck exposure over time

The dataset utilized in this study is comprised of 10 years of large truck crashes in work zone. It is unknown to the authors how large truck exposure has changed over time between the beginning and the ending years of the study. It was necessary to further expand the GORL model to better capture the true effects of time on the severity of the most injured person in a crash. Table 4 presents further time-of-day analysis conducted within the GORL model. Based on the hourly distribution of crashes within the dataset, several different categorizations of "time-of-day" variable were tested and compared using a Bayesian

Information Criterion (BIC) test, discussed in a later section, in order to arrive at the best-fit distribution of crashes over the different times of the day. The overall effect of the different years within the dataset was also tested to investigate how the severity of crashes involving large truck within a work zone might have changed over time.

Finally, a partially-segmented GORL model was developed using interactions between "time-of-day" variable and "year" index in order to better address the effects of time layered within years in a composite

TABLE 4 GORL-time-of-day interactions model results

	GORL	
Variable	**Latent Propensity**	**Threshold: injury | serious injury**
Roadway		
Functional class		
Principal arterial	0.609 (3.37)	-
Urban minor arterial	0.306 (1.66)	-
Geometric design		
Curbed		
Curb	-	0.253 (2.38)
Access-control		
No control	0.706 (5.31)	0.239 (2.31)
Traffic		
Speed limit (mph)		
< 45 mph	-0.663 (-4.54)	-
65 - 70 mph	0.267 (1.72)	-0.342 (-2.00)
Environmental		
Roadway surface condition		
Wet	-0.353 (-1.94)	0.133 (1.00)
Weather condition		
Adverse (rain, snow, fog, etc.)	0.230 (1.02)	-
Temporal		
Time-of-day		
Day (6:00 AM – 5:59 PM)	0.531 (3.02)	-0.514 (-2.52)
Year		
2003	0.297 (2.09)	-
2005	-	0.215 (1.36)
2006	-	1.046 (2.62)
2009	0.355 (2.24)	-
2010	0.314 (2.13)	-
Time-of-day and Year interactions		
Evening (6:00 PM – 11:59 PM) (year = 2006)	-	-0.771 (-1.79)

TABLE 4 continued

	GORL	
Variable	**Latent Propensity**	**Threshold: injury | serious injury**
Work Zone		
Work zone type		
Lane closure	-0.286 (-2.88)	-
Work zone location		
Transition area	-0.411 (-3.25)	-
Crash		
Number of vehicles		
Single-vehicle	-0.181 (-1.21)	-0.477 (-2.83)
Truck type		
1 Unit truck with trailer	0.230 (1.73)	-
Location		
On-bridge	0.245 (1.40)	-
Threshold coefficients		
No Injury | Possible Injury	0.6189 (5.84)	
Possible Injury | Serious Injury	0.3978 (4.78)	
Log-Likelihood at zero	-1,915.10	
Log-Likelihood at convergence	-1,839.13	
McFadden R^2	0.0397	
Number of observations	2,881	

way. The modified GORL model had positive coefficients for the years of 2003, 2009, and 2010 which essentially showed that crashes occurring during those years were more severe compared to crashes in other years within the dataset. On the other hand, years 2005 and 2006 indictors were associated with higher likelihood of sustaining “injury” but lower likelihood of “serious injury” relative to “no injury” outcomes in the event of a crash. Those results do not indicate sufficient evidence that work zone enforcement practices have changed to the better or worse over the years of the current study. Interactions of “time-of-day” variable and “year” index showed statistical significance for evening crashes in the year of 2006 and late night crashes in the year of 2008 in which both variables essentially showed that crashes occurring during those specific time periods in both years were more severe compared to other time periods. Based on the results of the modified GORL model, an overall conclusion of interactions of “time-of-day” variable with “year” index is that truck exposure did not change during the different times of the day across the years in this study.

7. Measures of Fit

The MNL and ORL models cannot be compared using the log-likelihood ratio test statistic because they are non-nested models. Also, when fitting a set of models, it is possible to increase the goodness-of-fit by adding more parameters but this may result in obtaining an over-fitted model. The Bayesian Information Criterion (BIC) controls for over-fitting in a model by introducing a penalty term in its calculation, which essentially grows with adding more parameters to the estimated model (Akaike, 1987; Schwarz, 1978). The model with the lowest BIC value is essentially the best-fit among all. As shown in the model comparison table, the MNL, ORL, and GORL had BIC values of 3912.21, 3876.35, and 3870.00 respectively indicating that GORL has the lowest BIC value and thus provides superior data fit among the three models for modeling crash severity data of work zones involving large trucks.

8. Elasticity Effects

The magnitude of the effects of the independent variables entering a statistical model on each severity outcome is not directly provided through the parameter values provided by the model. To be able to clearly understand the impacts of these variables, it is necessary to compute their corresponding elasticity effects. Elasticity effects can be interpreted as the percent effect a 1% change in a variable has on the severity outcome probability (Khorashadi et al., 2005). Elasticity calculations are not applicable to indicator variables; therefore average direct pseudo-elasticity was calculated (Chang and Mannering, 1999; Shankar and Mannering, 1996; Ulfarsson and Mannering, 2004). The pseudo-elasticity of a variable essentially represents the average percent change in the probability of an outcome category when the value of that variable is changed from 0 to 1. The elasticity analysis was undertaken only for the best model, *i.e.*, the GORL model.

8.1. Elasticity effects of GORL model

Aggregate level pseudo-elasticity effects of all the variables entered the GORL model were calculated and the results are shown in table 5. The numbers in the top row of Table 5 indicate that the elasticity effects of “Rural principal arterial” functional class for “No Injury”, “Injury”, and “Serious Injury” outcomes are –15.14%, 50.35%, and 76.93%, respectively. So, work zone crashes involving large trucks occurring on rural principal arterials are 15.14% less likely to result in “no injury” whereas 50.35% and 76.93% more likely to result in “Injury” and “Severe Injury” outcomes respectively compared to crashes on collectors, local system roads, and rural minor arterial. Other numbers in the table can be interpreted similarly.

Based on the elasticity effects, it can be seen that the key factors and conditions that increase the risk of severe outcomes of crashes involving large trucks in work zones are: daytime crashes, no control of access, higher speed limit, and rural principal arterials. Other variables such as urban principal arterial, one-unit truck with trailer, and single-vehicle also contribute to increased risk, but not as much as the variables identified earlier.

Variable effects have important implications for training and education for drivers, workers, and non-motorists. These implications could also be extended to the planning and design of a work zone area and the regulation and use of traffic control devices. In terms of training and education, the results suggest the importance of education to the drivers and training for work zone workers on daytime crash-developing situations in a work zone. It also suggests enforcing the use of highly reflective gears in work zones which increases the visibility of workers to the motorist. In terms of planning and design, the results suggest that roadways with no control of access require assigning additional traffic control devices. It is well known in the transportation field that traffic control devices in work zones mandated by the FHWA are the minimum to be used; therefore, extra traffic control measures may be warranted especially in areas with substantial large-truck traffic. Adding additional advisory and warning signs for non-motorists could effectively improve their alertness toward crash-developing situations.

Speed harmonization methods and increased presence of law enforcement officers are recommended for enforcing lower speeds especially on non-controlled access roadways where more conflict points are present. Rerouting truck-traffic away from work zones on rural principal arterials could decrease the

TABLE 5 Elasticity effects of GORL

Variable	No Injury	Injury	Serious Injury
	Mean	**Mean**	**Mean**
Roadway			
Functional class (base = other=collector, local system, rural minor arterial)			
Rural principal arterial	-15.14	50.35	76.93
Urban principal arterial	-8.61	34.72	47.19
Urban minor arterial	-5.11	16.67	22.90
Geometric design			
Curbed (base=no curb)			
Curb	0.00	11.96	-23.38
Access-control (base=full control, and partial control)			
No control	-12.87	29.83	128.40
Traffic			
Speed limit (mph) (base=speed limit 45-60 mph)			
< 35	14.69	-42.91	-50.08
35 - 40	7.87	-25.39	-30.84
65 - 70	-5.15	-8.37	83.46
Environmental			
Roadway surface condition (base=dry)			
Wet	7.03	-22.03	-27.10
Weather condition (base=clear)			
Adverse (rainy, snowy, foggy, etc.)	-4.73	15.29	20.95
Temporal			
Peak hours (base=off-peak)			
Peak	3.32	-10.27	-13.13
Time of the day (base=late night 12:00 AM - 6:00 AM)			
Day (6:00 AM - 6:00 PM)	-11.58	3.77	134.94
Work Zone			
Work zone type (base=ln shift/crossover, intermittent/moving work zone)			
Lane closure	5.23	-15.74	-19.84
Shoulder or median	-2.84	9.37	12.49
Work zone location (base=advanced signs, activity, termination, other areas)			
Transition	7.59	-23.57	-28.87
Crash			
Number of vehicles (base=multi-vehicle)			
Single-vehicle	3.92	-34.18	35.69
Truck type (base=bus, 2 axels 1 unit, other)			
1 unit with trailer	-7.62	24.50	34.52
Location			
On-bridge (base=not on-bridge)	-5.21	16.79	23.08

severity of a crash on this type of a functional class. Splitting truck traffic from other traffic will reduce conflicts in a work zone as well as give more space to non-heavy truck traffic for more flexible maneuvering to avoid possible crash situations.

In terms of regulation of traffic, the results suggest to extend lower speed limits prior to entering work zone areas, which will allow more time for drivers to recognize the setup of the specific work zone being approached. It is essential to post traffic control signs that can communicate to vehicle drivers and non-motorists of sharing the roadway with large-truck traffic.

9. Conclusions

Safety literature focusing on work zone safety of large trucks is sparse. This research effort aims to fill this gap in the literature by undertaking an extensive empirical analysis of large truck crashes in work zones by pooling together 10 years of crash databases in the State of Minnesota. The empirical analysis employs statistical models that encompass recent advances in the econometric literature. Specifically, both unordered and ordered modeling methods were deployed and the best modeling method for the current empirical context was chosen. To our knowledge, this is first such comparison of a comprehensive set of discrete choice models in the context of work zone safety.

A wide array of explanatory variables characterizing the crash, roadway, and work zone conditions were considered in the model estimation process. All models were gradually fine-tuned by removing statistically insignificant variables until the best-fit specification was obtained. In the unordered framework, the multinomial logit (MNL) and nested logit (NL) models were estimated. The NL model was used to test the validity of the IIA assumption in MNL model given the intrinsic ordinal nature of injury severity data being modeled. In the ordered response framework, simple ordered response logit (ORL) and generalized ordered response logit (GORL) models that explicitly recognize the ordinal nature of severity outcomes were estimated. The GORL model is a generalized version of the standard ORL model that relaxes the fixed thresholds assumption of the ORL thus providing additional flexibility. The performance of different models developed in this study was compared using Bayesian Information Criterion (BIC) test statistic. Among all the different models estimated in this study, the GORL model was found to offer the best-fit as indicated by its lower BIC value compared to other models. Lastly, going beyond simple parameter estimates, elasticity effects were computed to quantify the magnitude of impact of different exogenous factors considered in the study.

There are important empirical findings in the current study. The GORL model elasticity effects indicate that the most important factors/conditions that contribute to higher severity outcomes in the event of a crash are: daytime crashes, no control of access, higher speed limits, and crashes on rural principal arterials. Other variables such as urban principal arterial, one-unit truck with trailer, and single-vehicle also contribute to higher risk, but not as much as the variables identified earlier. With regards to potential improvements to this study, the authors used 10 years of crash data from the State of MN due to work zone data availability. So, the study findings may not be extended to all work zones in the nation given that unique conditions specific to locations in MN may have influenced the analysis. Future research studies using combined datasets across multiple states will provide more evidence and confidence in the study findings. Also, bigger datasets allow segmentation of single and multi-vehicle crashes (*i.e.*, single truck crashes versus truck and car collisions) to check if there are significant differences in factors affecting severity of these two types of crashes. Another avenue for future research is exploring the endogeneity of work zone by including both work and non-work zone crashes in the analysis. This is important because injury severity outcomes at a work zone site can be more (or less) severe because of unobserved factors that caused the site to be a work zone site. Simultaneous modeling methods that jointly analyze crash occurrence at a work zone and severity conditional on crash occurrence in a work zone will enable unbiased estimation of model parameters (Eluru and Bhat, 2007; Kim and Washington, 2006). Future research including work zone-specific data such as modified lane, shoulder, and median widths, lengths of areas composing a work zone, and specific work zone speed limits could be beneficial. Also, in this study, we focused only on crash severity defined as the severity level of the most severely injured person in the crash. However, future studies can conduct occupant-level analysis that considers all people involved in the crash. This is important to obtain better insights into the relative profile of different occupant risk propensities and their determinants.

References

Akaike, H., 1987. Factor analysis and AIC. Psychometrika 52, 317–332. doi:10.1007/BF02294359

Anowar, S., Yasmin, S., Tay, R., 2014. Factors Influencing the Severity of Intersection Crashes in Bangladesh. Asian Transp. Stud. 3, 143–154. doi:10.11175/eastsats.3.143

Chang, L.-Y., Chien, J.-T., 2013. Analysis of driver injury severity in truck-involved accidents using a non-parametric classification tree model. Saf. Sci. 51, 17–22. doi:10.1016/j.ssci.2012.06.017

Chang, L.-Y., Mannering, F., 1999. Analysis of injury severity and vehicle occupancy in truck- and non-truck-involved accidents. Accid. Anal. Prev. 31, 579–592. doi:10.1016/S0001-4575(99)00014-7

Chen, F., Chen, S., 2011. Injury severities of truck drivers in single- and multi-vehicle accidents on rural highways. Accid. Anal. Prev. 43, 1677–1688. doi:10.1016/j.aap.2011.03.026

Chen, P., Shen, Q., 2016. Built environment effects on cyclist injury severity in automobile-involved bicycle crashes. Accid. Anal. Prev. 86, 239–246. doi:10.1016/j.aap.2015.11.002

Dong, C., Richards, S.H., Huang, B., Jiang, X., 2015. Identifying the factors contributing to the severity of truck-involved crashes. Int. J. Inj. Contr. Saf. Promot. 22, 116–126. doi:10.1080/17457300.2013.844713

Duncan, C., Khattak, A., Council, F., 1998. Applying the Ordered Probit Model to Injury Severity in Truck-Passenger Car Rear-End Collisions. Transp. Res. Rec. J. Transp. Res. Board 1635, 63–71. doi:10.3141/1635-09

Eluru, N., 2013. Evaluating alternate discrete choice frameworks for modeling ordinal discrete variables. Accid. Anal. Prev. 55, 1–11. doi:10.1016/j.aap.2013.02.012

Eluru, N., Bhat, C.R., 2007. A joint econometric analysis of seat belt use and crash-related injury severity. Accid. Anal. Prev. 39, 1037–1049. doi:10.1016/j.aap.2007.02.001

Eluru, N., Bhat, C.R., Hensher, D.A., 2008. A mixed generalized ordered response model for examining pedestrian and bicyclist injury severity level in traffic crashes. Accid. Anal. Prev. 40, 1033–1054. doi:10.1016/j.aap.2007.11.010

FHWA, 2014. Large Truck and Bus Crash Facts 2012 [WWW Document]. Fed. Mot. Carr. Saf. Adm. URL http://www.fmcsa.dot.gov/safety/data-and-statistics/large-truck-and-bus-crash-facts-2012 (accessed 7.21.15).

FHWA, 2012. Highway Statistics 2012, Highway Statistics Series. FHWA - Office of Highway Policy Information.

FHWA, 2010. Work Zone Facts and Statistics [WWW Document]. URL http://www.ops.fhwa.dot.gov/wz/resources/facts_stats/injuries_fatalities.htm (accessed 7.21.15).

Islam, Hernandez, 2013. Large Truck–Involved Crashes: Exploratory Injury Severity Analysis. J. Transp. Eng. 139, 596–604. doi:10.1061/(ASCE)TE.1943-5436.0000539

Islam, M., Hernandez, S., 2013. Modeling Injury Outcomes of Crashes Involving Heavy Vehicles on Texas Highways. Transp. Res. Rec. J. Transp. Res. Board 2388, 28–36. doi:10.3141/2388-05

Khorashadi, A., Niemeier, D., Shankar, V., Mannering, F., 2005. Differences in rural and urban driver-injury severities in accidents involving large-trucks: An exploratory analysis. Accid. Anal. Prev. 37, 910–921. doi:10.1016/j.aap.2005.04.009

Kim, D.-G., Washington, S., 2006. The significance of endogeneity problems in crash models: An examination of left-turn lanes in intersection crash models. Accid. Anal. Prev. 38, 1094–1100. doi:10.1016/j.aap.2006.04.017

Lemp, J.D., Kockelman, K.M., Unnikrishnan, A., 2011. Analysis of large truck crash severity using heteroskedastic ordered probit models. Accid. Anal. Prev. 43, 370–380. doi:10.1016/j.aap.2010.09.006

Li, Y., Bai, Y., 2009. Highway Work Zone Risk Factors and Their Impact on Crash Severity. J. Transp. Eng. 135, 694–701. doi:10.1061/(ASCE)TE.1943-5436.0000055

Li, Y., Bai, Y., 2008. Development of crash-severity-index models for the measurement of work zone risk levels. Accid. Anal. Prev. 40, 1724–1731. doi:10.1016/j.aap.2008.06.012

Manski, C.F., McFadden, D. (Eds.), 1981. Structural Analysis of Discrete Data with Econometric Applications. The MIT Press, Cambridge, Mass.

McKelvey, R.D., Zavoina, W., 1975. A statistical model for the analysis of ordinal level dependent variables. J. Math. Sociol. 4, 103–120. doi:10.1080/0022250X.1975.9989847

Mooradian, J., Ivan, J.N., Ravishanker, N., Hu, S., 2013. Analysis of driver and passenger crash injury severity using partial proportional odds models. Accid. Anal. Prev. 58, 53–58. doi:10.1016/j.aap.2013.04.022

Pahukula, J., Hernandez, S., Unnikrishnan, A., 2015. A time of day analysis of crashes involving large trucks in urban areas. Accid. Anal. Prev. 75, 155–163. doi:10.1016/j.aap.2014.11.021

Qi, Y., Srinivasan, R., Teng, H., Baker, R., 2013. Analysis of the frequency and severity of rear-end crashes in work zones. Traffic Inj. Prev. 14, 61–72. doi:10.1080/15389588.2012.675109

Romo, A., Hernandez, S., Cheu, R.L., 2014. Identifying Precrash Factors for Cars and Trucks on Interstate Highways: Mixed Logit Model Approach. J. Transp. Eng. 140, 04013016. doi:10.1061/(ASCE)TE.1943-5436.0000621

Schwarz, G., 1978. Estimating the Dimension of a Model. Ann. Stat. 6, 461–464. doi:10.1214/aos/1176344136

Shankar, V., Mannering, F., 1996. An exploratory multinomial logit analysis of single-vehicle motorcycle accident severity. J. Safety Res. 27, 183–194. doi:10.1016/0022-4375(96)00010-2

Shankar, V., Mannering, F., Barfield, W., 1996. Statistical analysis of accident severity on rural freeways. Accid. Anal. Prev. 28, 391–401. doi:10.1016/0001-4575(96)00009-7

Ulfarsson, G.F., Mannering, F.L., 2004. Differences in male and female injury severities in sport-utility vehicle, minivan, pickup and passenger car accidents. Accid. Anal. Prev. 36, 135–147. doi:10.1016/S0001-4575(02)00135-5

U.S. Department of Transportation, 2014. Traffic Safety Facts 2012 (No. DOT HS 812 032), Traffic Safety Fact Annual Report. National Highway Traffic Safety Administration (NHTSA).

Wang, H.L., Shi, Z.K., 2013. Hierarchical Regression Model for Truck Collision Severity Analysis. Adv. Mater. Res. 671-674, 2889–2892. doi:10.4028/www.scientific.net/AMR.671-674.2889

Wang, Z., Lu, J., Wang, Q., Lu, L., Zhang, Z., 2010. Modeling Injury Severity in Work Zones using Ordered PROBIT Regression, in: ICCTP 2010. American Society of Civil Engineers, pp. 1058–1067.

Wong, J.M., Arico, M.C., Ravani, B., 2011. Factors Influencing Injury Severity to Highway Workers in Work Zone Intrusion Accidents. Traffic Inj. Prev. 12, 31–38. doi:10.1080/15389588.2010.525569

Yasmin, S., Eluru, N., 2013. Evaluating alternate discrete outcome frameworks for modeling crash injury severity. Accid. Anal. Prev. 59, 506–521. doi:10.1016/j.aap.2013.06.040

Yasmin, S., Eluru, N., Ukkusuri, S.V., 2014. Alternative Ordered Response Frameworks for Examining Pedestrian Injury Severity in New York City. J. Transp. Saf. Secur. 6, 275–300. doi:10.1080/19439962.2013.839590

Ye, F., Lord, D., 2014. Comparing three commonly used crash severity models on sample size requirements: Multinomial logit, ordered probit and mixed logit models. Anal. Methods Accid. Res. 1, 72–85. doi:10.1016/j.amar.2013.03.001

Zhu, X., Srinivasan, S., 2011a. A comprehensive analysis of factors influencing the injury severity of large-truck crashes. Accid. Anal. Prev. 43, 49–57. doi:10.1016/j.aap.2010.07.007

Zhu, X., Srinivasan, S., 2011b. Modeling occupant-level injury severity: An application to large-truck crashes. Accid. Anal. Prev. 43, 1427–1437. doi:10.1016/j.aap.2011.02.021