Accident Analysis and Prevention 102 (2017) 93–100

Contents lists available at ScienceDirect

Accident Analysis and Prevention

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

An empirical analysis of run-off-road injury severity crashes involving large trucks

Nabeel Saleem Saad Al-Bdairi, Salvador Hernandez *

School of Civil and Construction Engineering, Oregon State University, 101 Kearney Hall, Corvallis, OR 97331-3212, United States

ARTICLE INFO

Article history:
Received 8 November 2016
Received in revised form 22 February 2017
Accepted 24 February 2017

Keywords:
Large-truck safety
Run-off-road crashes
Ordered random parameter probit model
Injury severity
Unobserved heterogeneity

ABSTRACT

In recent years, there has been an increasing interest in understanding the contributory factors to run-off-road (ROR) crashes in the US, especially those where large trucks are involved. Although there have been several efforts to understand large-truck crashes, the relationship between crash factors, crash severity, and ROR crashes is not clearly understood. The intent of this research is to develop statistical models that provide additional insight into the effects that various contributory factors related to the person (driver), vehicle, crash, roadway, and environment have on ROR injury severity. An ordered random parameter probit was estimated to predict the likelihood of three injury severity categories using Oregon crash data: severe, minor, and no injury. The modeling approach accounts for unobserved heterogeneity (i.e., unobserved factors). The results showed that five parameter estimates were found to be random and normally distributed, and varied across ROR crash observations. These were factors related to crashes that occurred between January and April, raised median type, loss of control of a vehicle, the indicator variable of speed not involved, and the indicator variable of two vehicles or more involved in the crashes. In contrast, eight variables were found to be fixed across ROR observations. Looking more closely at the fixed parameter results, large-truck drivers who are not licensed in Oregon have a higher probability of experiencing no injury ROR crash outcomes, and human related factor, fatigue, increases the probability of minor injury. The modeling framework presented in this work offers a flexible methodology to analyze ROR crashes involving large trucks while accounting for unobserved heterogeneity. This information can aid safety planners and the trucking industry in identifying appropriate countermeasures to help mitigate the number and severity of large-truck ROR crashes.

1. Introduction

In recent years, there has been an increasing interest in understanding the contributory factors to run-off-road (ROR) crashes in the US, especially those where large trucks are involved (Davis et al., 2006; Lee and Mannering, 2002; McLaughlin et al., 2009; Peng and Boyle, 2012; Roy and Dissanayake, 2011). One reason for this is that in 2010, approximately 57% of all fatal crashes were ROR crashes, whereas nonfatal crashes accounted for 16% (Blincoe et al., 2015). Accordingly, those crashes led to roughly $64 billion in economic costs and $298 billion in comprehensive costs, accounting for 27% of all economic costs and 36% of all societal harm (Blincoe et al., 2015). Although, statistically, the number of large-truck-involved crashes has decreased over the past two decades, there is still a higher fatal crash involvement rate per 100 million vehicle miles traveled compared to passenger cars (1.34 versus 1.08 for the year 2014) (Federal Motor Carrier Safety Administration, 2016). As a result, several state agencies have today developed and/or have adopted mitigation programs to reduce the number and severity of these crashes. For example, in Oregon, where nearly 66% of all fatal crashes were due to ROR crashes in 2010, the Oregon Department of Transportation (ODOT) partnered with the Federal Highway Administration (FHWA) to implement appropriate and low-cost countermeasures with the goal of reducing the number of ROR fatalities by 20%. However, the implemented countermeasures focused primarily on reducing ROR crashes for passenger cars, with little focus on large trucks (gross vehicle weight rating [GVWR] greater than 10,000 pounds). With this in mind, there is a clear need for continued research into identifying and/or developing cost-effective countermeasures to reduce the number and severity of ROR crashes involving large trucks.

Various methodological models have been used in analyzing severity and frequency of crashes. The selection of an appropri-

* Corresponding author.
E-mail addresses: albdairn@oregonstate.edu (N.S.S. Al-Bdairi), sal.hernandez@oregonstate.edu (S. Hernandez).

http://dx.doi.org/10.1016/j.aap.2017.02.024
0001-4575/© 2017 Published by Elsevier Ltd.

ate statistical model depends primarily on the nature of the crash data. In this study, the injury severity sustained by trucks' drivers are the main interest. Therefore, only previous works that examined the crash severity will be reviewed. The current study has aimed to examine the impact of contributory factors on the run-off-the-road (ROR) crashes that involved large trucks. In terms of the risk factors, several studies have been conducted to investigate the effect of some possible factors on ROR crash severity. In general, these factors can be summarized into three main groups, human factors, highways' geometric design and environmental factors, and roadside factors. In terms of human factors, Davis et al. (2006) reviewed the previous works that studied the effect of speed on ROR crashes on rural two-lane highways. They collected data from Australia and Minnesota for their study and used Bayesian relative risk regression. In their study, they found that high speed was associated with higher fatality risk. McGinnis et al. (2001) used Fatality Analysis Reporting System (FARS) data for years 1975,1980, 1985, 1990, 1996, and 1997 to investigate the contributory factors that might affect ROR crashes. They found that around half of ROR crashes occurred due to intoxicated drivers, particularly male drivers whose age between 20 and 39 years. They also found that severity of ROR crashes that involved male drivers were higher than ROR crashes that involved female drivers.

Turning to statistical approaches, several models have been used to determine the relationship between the potential contributory factors and ROR crash severity. Liu and Subramanian (2009) conducted a univariate analysis with Chi-square tests and logistic regression to analyze FARS crash data for the period from 1991 to 2007. In their study, they attempted to capture the effect of various factors on ROR crashes such as roadway and environmental factors, driver characteristics, and traffic-related factors. They stated that some variables were statistically significant and affected ROR crashes such as the presence of horizontal curves on the roadway, alcohol impairment, number of lanes, inclement weather, and driver age. Roy and Dissanayake (2011) developed a Bayesian statistical approach to compare ROR with non-ROR crashes in Kansas by using crash data for crashes that occurred in the period between 1999 and 2008. They found some variables were highly associated with ROR crashes rather than non-ROR crashes. These variables included road surface condition (i.e., wet and icy surfaces), time of the day (i.e., nighttime), rural area, inclement weather conditions, horizontal curve sections, higher speed, and fatigue and drowsiness. Dissanayake (2003) conducted a study to identify the contributory factors that affect the severity of ROR crashes involving young drivers with age from 16 to 25 years. In this study, Dissanayake obtained a crash data from Florida traffic crash database. He categorized the injury severity into five injury outcomes and then developed four sequential binary logistic regression models. He concluded that some factors were highly influencing the severity of young drivers involved in ROR crashes such as gender, lighting condition, area type, and roadway alignment.

The majority of the aforementioned studies primarily focused on analyzing the injury severity of passenger cars involved ROR crashes. However, works that study the injury severity of drivers involved in large trucks ROR crashes are sparse. Some studies in recent years have specifically studied large-truck-involved crashes from various perspectives. Some of this work has dealt with understanding the risk and human-related factors of ROR crashes caused by speed, driver characteristics, driving under the influence of alcohol and/or drug impairments, fatigue or drowsiness, roadway characteristics, vehicle characteristics, and environmental factors (Aram, 2010; Compton and Berning, 2009; LeRoy et al., 2008; McGinnis et al., 2001; Neuman et al., 2003; Peng and Boyle, 2012; NHTSA, 2012). Other studies have focused on identifying the contributory factors to large-truck-involved crashes through econometric and statistical models. In those studies, ROR crashes are represented as an indicator variable for crashes related to urban settings, rural versus urban, time of day, manner of collision, and vehicle type (Cerwick et al., 2014; Chang and Mannering, 1999; Chen and Chen, 2011; Duncan et al., 1998; Islam and Hernandez, 2013a,b, 2015; Islam et al., 2014; Khorashadi et al., 2005; Lemp et al., 2011; Pahukula et al., 2015; Romo et al., 2014).

Although there have been several efforts to understand large-truck crashes, the relationship between contributory factors and severity of ROR crashes is not clearly understood. One reason for this stems from the lack of detailed crash data to capture the complex interactions of multiple factors under a single framework for ROR crashes. Therefore, the purpose of this research is to develop statistical models that provide additional insight into the effects that various contributory factors related to the person (driver), vehicle, crash, roadway, and environment have on ROR injury severity. This is done by analyzing the Oregon Statewide Crash Data System, which is an extensive database collected and maintained by ODOT. The findings of this study can provide information that can aid safety planners and the trucking industry in identifying appropriate countermeasures to help mitigate the number and severity of large-truck ROR crashes. To the best of the authors' knowledge, these are the first attempts at developing these types of models for ROR crashes.

The rest of the paper is organized as follows. In Section 2, the crash data used in the analysis and their descriptive statistics are described. Section 3 presents details of the proposed econometric modeling framework. In Section 4, estimation results along with discussions are presented. Section 5 provides conclusions and suggestions for future research.

2. Data description

This study utilizes data collected from the Oregon Statewide Crash Data System provided by ODOT. The data obtained represents seven years of large-truck-involved crashes, from 2007 to 2013; large-truck-involved crashes for the seven-year period comprised 13,364 records. However, since ROR crashes are the main interest of this study, only crashes belonging to this category are considered, bringing the sample size down to roughly 2486 observations (data filtered by ROR indicator). Each ROR observation represents the maximum level of injury severity sustained by the driver following the National Safety Council (NSC) injury severity scale, KABCO. The KABCO injury severity scale characteristically consists of five injury categories: fatality (K), incapacitating (A), non-incapacitating (B), possible injuries (C), and non-injury (O) or property damage only (PDO). For this study, any recorded incidents that showed an injury severity of "not reported" or "unknown" were rejected because the severity of those injuries could not be satisfactorily determined. As was the case with other studies (Anarkooli and Hosseinlou, 2016; Haleem and Abdel-Aty, 2010; Haleem and Gan, 2013; Pahukula et al., 2015; Quddus et al., 2002), because of low data observations for the higher injury severity outcomes, the full KABCO scale was reduced to three injury categories. These categories are severe injury (KA- fatal and incapacitating), minor injury (BC- non-incapacitating and possible injury), and no injury (O- property damage only or PDO).

Turning to the data, overall severe injury, minor injury, and no injury accounted for 2.6% (N = 65), 24.6% (N = 612), and 72.8% (N = 1809), respectively. Table 1 illustrates the descriptive statistics of key variables for large-truck-ROR crash severity. These variables were selected according to their statistical significance and minimal correlation.

In terms of the driver-related factors, injury statistics, shown in Table 1, show that 2.4% of truck drivers who were involved in

Table 1
Frequency and percentage distribution of driver injury.

Variable	Severe injury	Minor injury	No injury	Total
Month of the year (1 if crash occurred between January and April, 0 otherwise)	13 (1.8%)	148 (20.7%)	555 (77.5%)	716
Median type (1 for raised median, 0 otherwise)	6 (1.3%)	73 (16.9%)	354 (81.8%)	433
Roadway surface condition (1 for dry, 0 otherwise)	50 (3.6%)	387 (27.5%)	969 (68.9%)	1406
Driver license status (1 for a license from other states or countries, 0 otherwise)	20 (1.9%)	242 (22.6%)	810 (75.5%)	1072
Vehicle maneuver just before impending crash (1 if going straight, 0 otherwise)	62 (3.0%)	578 (28.1%)	1417 (68.9%)	2057
Participant level action (1 for lost control of a vehicle, 0 otherwise)	32 (3.5%)	291 (32.2%)	582 (64.3%)	905
Participant level safety equipment use (1 if seatbelt was fastened, 0 otherwise)	33 (1.6%)	511 (25.7%)	1445 (72.7%)	1989
Crash type (1 for overturn, 0 otherwise)	8 (2.0%)	182 (45.5%)	210 (52.5%)	400
Alcohol not a factor (1 for no alcohol involved, 0 otherwise)	59 (2.4%)	605 (24.7%)	1789 (72.9%)	2453
Speed not a factor (1 for non-speed involved, 0 otherwise)	62 (2.5%)	589 (24.2%)	1784 (73.3%)	2435
Roadway characteristics (1 for horizontal curve, 0 otherwise)	25 (3.5%)	251 (35.5%)	431 (61.0%)	707
Participant level action (1 for fatigue, 0 otherwise)	0 (0.0%)	48 (43.6%)	62 (56.4%)	110
Number of vehicles involved in the crash (1 if two or more vehicles, 0 otherwise)	6 (1.0%)	64 (10.6%)	532 (88.4%)	602

ROR crashes sustained severe injury when they were not under the influence of alcohol at the time of the crash, whereas 72.9% of those drivers sustained no injury. Moreover, 75.5% of drivers whose driver license were not issued in Oregon were experienced no injury, while drivers who were sustained severe injury were 1.9%. The potential reason might be related to driver behavior, probably because their driving is more cautious since they are unfamiliar with Oregon highways. Furthermore, fatigued drivers were found to be more likely to sustain minor- and no-injury since injury statistics show that those injury categories are 43.6% and 56.4%, respectively. In addition, losing control of a vehicle constituted as highly as effect than aforementioned driver related factors because 3.5% of ROR crashes that occurred due to losing control of vehicles caused severe injury for drivers.

With regard to roadway and environmental factors, ROR crashes that occurred on horizontal curves and on dry roadway surfaces were associated with high probability of sustaining severe injury. Table 1 depicts that those crashes have slightly similar impact on drivers' injury since 3.5% of ROR crashes on horizontal curves and 3.6% of ROR crashes on dry roadway surfaces caused severe injury for drivers. Moreover, it is quite interesting to note that 77.5% of ROR crashes that occurred between January and April were associated with no injury outcome, whereas minor and severe injury outcomes were 20.7% and 1.8%, respectively. One potential explanation is Oregonian drivers might be accustomed to the prevalent adverse and inclement weather conditions in the aforementioned period. For more details regarding percentage distribution of ROR crashes for each injury category (see Table 1).

It should be noted that some variables, such as the type of shoulder, shoulder width, lane width, and the number of lanes, were not considered in the analysis. This is because the information regarding those variables was unavailable or the crash dataset had a large proportion of missing data for the variables not considered. Furthermore, the 95% confidence level was used to gauge the statistical significance of the selected variables (see Results and Discussion section). Consequently, statistically insignificant variables were not considered.

3. Methodological approach

With regard to methodological approaches, many applications of statistical modeling methods have been applied in recent years to a variety of injury severity analysis scenarios. Mannering and Bhat (2014) provide a complete review of these applications. However, most of these studies have focused primarily on crash data related to a passenger car or all traffic mixes in a single modeling framework. With regard to large-truck-involved crashes, studies on modeling injury severity analysis are sparse, and a variety of statistical modeling frameworks are used, depending on the definition of the variables of interest (e.g., how injury severity is defined) (Cerwick et al., 2014; Chen and Chen, 2011; Duncan et al., 1998; Islam and Hernandez, 2013a,b; Islam et al., 2014; Khorashadi et al., 2005; Lemp et al., 2011; Pahukula et al., 2015). In this work, three injury categories are used, as previously defined, to model injury severity of ROR crashes: severe injury, minor injury, and no injury.

With this in mind, to investigate the relationship between the injury severity of ROR crashes and the possible contributory factors, an ordered probit modeling framework is considered. Traditionally, ordered probit models have been used to model and account for the ordinal nature of injury severity data. However, it has been shown that the traditional (or fixed parameter) ordered probit model is susceptible to underreporting of crash injury severity. Thus, the fixed parameter ordered probit framework imposes restrictions on the impacts of explanatory variables that are assumed to be the same across individual injury observations, whereas in ordered random parameter probit model explanatory variables are assumed to be varied across the injury observations (Eluru et al., 2008; Eluru and Yasmin, 2015; Russo et al., 2014; Savolainen et al., 2011). These drawbacks in ordered probit model with fixed parameters can lead to inconsistent (i.e., incorrect) estimates of the effects of variables on injury severity (Abdel-Aty, 2003; Abdel-Aty and Keller, 2005; Anarkooli and Hosseinlou, 2016; Haleem and Abdel-Aty, 2010; Obeng, 2011). Therefore, two approaches have been proposed to overcome the restrictions in the traditional ordered probit model (Eluru and Yasmin, 2015). These approaches are either allowing thresholds to be random or the effect of explanatory variables treated as random (Eluru and Yasmin, 2015). Some previous works have been conducted in context of traffic injury by following the first approach proposed by Eluru and Yasmin (2015) (i.e., treating thresholds as random) to relax the aforementioned restrictions in traditional ordered probit models. For instance, Eluru et al. (2008), Srinivasan (2002), and Yasmin and Eluru (2013) used mixed generalized ordered logit (MGOL) to account for unobserved heterogeneity in the effect of explanatory variables on injury severity levels in both the latent injury risk propensity function and the threshold functions.

On the other hand, Eluru et al. (2008), Eluru and Yasmin (2015), Hensher et al. (2015), Russo et al. (2014), and Savolainen et al. (2011) argue that to overcome the aforementioned drawbacks of the ordered probit model with fixed parameter, extending the ordered probit model to an ordered random parameter probit model that accounts for unobserved heterogeneity (also referring to unobserved factors), can account for the above drawbacks. In this paper, the impact of explanatory variables has been treated as random to overcome the restrictions of traditional ordered probit model.

To illustrate the superiority of the ordered random parameter probit model, this model will be compared with the fixed parameter ordered probit model.

3.1. Ordered random parameter probit model

The ordered random parameter probit model is formulated by specifying an unobserved variable, z, as a linear function of a vector of explanatory variables $\boldsymbol{X}$ and the associated vector of estimable parameters $\boldsymbol{\beta}$ (e.g., person [driver], vehicle, crash, roadway, and environment), along with an error term or a disturbance term ε, which is assumed to be independently randomly distributed, with a mean of 0 and a variance of 1 (Eluru et al., 2008; Hensher et al., 2015; Washington et al., 2011). The unobserved variable, z, can be represented as illustrated in Eq. (1):

$$z = \boldsymbol{\beta X} + \varepsilon$$

Then, by using Eq. (1) for each observation, ordinal injury data y can be defined as shown in Eq. (2)

$$\begin{aligned} &y = 1 \quad \text{if} z \le \mu_0 \\ &y = 2 \quad \text{if} \mu_0 < z \le \mu_1 \\ &y = 3 \quad \text{if} \mu_1 < z \le \mu_2 \\ &y = \ldots \\ &y = J \quad \text{if} z \ge \mu_{J-1} \end{aligned}$$

where μ equals estimable parameters or thresholds that define the ordinal injury data y. In general, these thresholds are estimated jointly with model estimable parameters $\boldsymbol{\beta}$, which corresponds to integer ordering, and where J represents the highest integer ordered response (in this work, that response is no injury).

Next, to estimate the probabilities of J, the specific ordered response for each ROR crash observation, the error term (ε) is assumed to be normally distributed, with mean and variance equal to 0 and 1, respectively. The ordered selection probabilities are illustrated in Eq. (3):

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

where the highest (y = 3) represents no injury and the lowest (y = 1) represents severe injury.

As an attempt to account for unobserved heterogeneity and to allow a variable to have various effects across the observations, an ordered probit model with random parameters is applied. The simulated maximum likelihood estimation procedure was established by Greene to use random parameters in the ordered probit models, as illustrated in Eq. (4) (Greene, 2007).

$$\beta_i = \boldsymbol{\beta} + u_i \tag{4}$$

where u_i is a randomly distributed term (for example, a normally distributed term, with mean 0 and variance σ^2). Estimation of the ordered random parameter probit model is accomplished through the use of the Halton sequence approach (Anastasopoulos and Mannering, 2009; Bhat, 2003; Train, 1998). In this study, 200 Halton draws were used, and several random parameter distributions were examined, such as the normal, lognormal, triangular, and uniform distributions (Anastasopoulos and Mannering, 2009). However, only the normal distribution produced statistically significant results.

To investigate the impact of a particular variable on the injury outcomes, the marginal effects are used. The marginal effects represent the change in the probability of a particular injury outcome due to one unit change in an explanatory variable while holding all other variables constant. Moreover, the marginal effects are commonly used along with an ordered probit model to help interpret the interior injury outcomes or thresholds. The marginal effects corresponding to the probability of each category can be estimated as illustrated in Eq. (5) (Washington et al., 2011):

$$\begin{aligned} &\frac{\partial P(y=1)}{\partial X} = -\emptyset(-\boldsymbol{\beta X})\boldsymbol{\beta}' \\ &\frac{\partial P(y=2)}{\partial X} = [\emptyset(\mu_0 - \boldsymbol{\beta X}) - \emptyset(\mu_1 - \boldsymbol{\beta X})]\boldsymbol{\beta}' \\ &\frac{\partial P(y=3)}{\partial X} = [\emptyset(\mu_1 - \boldsymbol{\beta X}) - \emptyset(\mu_2 - \boldsymbol{\beta X})]\boldsymbol{\beta}' \\ &\frac{\partial P(y=J)}{\partial X} = -\emptyset(\mu_{J-1} - \boldsymbol{\beta X})\boldsymbol{\beta}' \end{aligned} \tag{5}$$

4. Results and discussion

Using simulation-based maximum likelihood and maximum likelihood methods with 200 Halton draws, an ordered probit model with random parameters was estimated. The purpose for using 200 Halton draws was to get a precise and accurate estimate of the random parameters (Bhat, 2003). With regard to the distribution of the random parameters in this analysis, the uniform, triangular, lognormal, and normal distributions were tested, and the normal distribution was the only distribution that resulted in statistically significant estimates for the random parameters. The econometric software NLOGIT 5.0 was used to analyze the data and fit the ordered probit model with fixed parameters and random parameters. The estimated results for the ordered probit model with fixed and random parameters along with the marginal effects are shown in Tables 2 and 3 respectively.

The estimated results presented in Table 2 illustrate that only five variables were found to be random and statistically significant (the standard deviation was statistically significant or different from zero), whereas the rest of other variables, along with the constant term, were found to be fixed variables and did not vary across the observations. The random variables were crashes that occurred between January and April, raised median type, loss of control of a vehicle, the indicator variable of speed not involved, and the indicator variable of two vehicles or more involved in the crashes. All the random variables were normally distributed. The criterion that distinguishes between the fixed- and random- parameters in a model is the statistical significance of the standard deviation corresponding to each variable. In other words, if the standard deviation corresponding to a particular variable is significant, and different from zero, that variable will be a random variable, and vice versa (Agbelie, 2014).

To test the null hypothesis that there is no statistical difference between the ordered probit model with fixed parameters and ordered probit model with random parameters in representing the provided data, a likelihood ratio test was performed. According to Washington et al. (2011), the likelihood ratio test is written as illustrated in Eq. (6):

$$\chi^2 = -2\left[LL_{\beta Fixed} - LL_{\beta Random}\right]$$

$$\chi^2 = -2\,[-1504.41 - (-1498.32)] = 12.18$$

Together, the Chi-square statistic value of 12.18 and corresponding degrees of freedom, 5 (number of random parameters), give more than 96.0% confidence level that the ordered probit model with random parameters is superior to the ordered probit model with fixed parameters. These results indicate that the null hypothesis should be rejected. Moreover, Table 2 shows that log-likelihood at convergence for the ordered probit model with random parameters is significantly better than the ordered probit model with

Table 2
Estimated Results for the Ordered Probit Model with Fixed and Random Parameters.

Variable	Fixed Parameter Model		Random Parameter Model	
	Coefficient	t-stat	Coefficient	t-stat
Constant	1.772	5.90	1.863	6.35
Month of the year (1 if crash occurred between January and April, 0 otherwise) *(standard deviation)*	0.122	1.90	0.161 (0.268)	2.35 (4.72)
Median type (1 for raised median, 0 otherwise) *(standard deviation)*	0.262	3.23	0.361 (0.448)	3.99 (5.44)
Roadway surface condition (1 for dry, 0 otherwise)	−0.214	−3.56	−0.239	−3.79
Driver license status (1 for a license from other states or countries, 0 otherwise)	0.162	2.83	0.178	2.95
Vehicle maneuver just before impending crash (1 if going straight, 0 otherwise)	−0.755	−7.71	−0.873	−8.24
Participant level action (1 for lost control of a vehicle, 0 otherwise) *(standard deviation)*	−0.242	−3.97	−0.239 (0.141)	−3.91 (3.22)
Participant level safety equipment use (1 if seatbelt was fastened, 0 otherwise)	0.223	3.18	0.276	3.83
Crash type (1 for overturn, 0 otherwise)	−0.310	−4.45	−0.322	−4.25
Alcohol not a factor (1 for no alcohol involved, 0 otherwise)	0.592	2.75	0.682	3.35
Speed not a factor (1 for non-speed involved, 0 otherwise) *(standard deviation)*	0.342	2.03	0.335 (0.133)	1.94 (4.58)
Roadway characteristics (1 for horizontal curve, 0 otherwise)	−0.207	−3.47	−0.232	−3.74
Participant level action (1 for fatigue, 0 otherwise)	−0.270	−2.16	−0.254	−1.76
Number of vehicles involved in the crash (1 if two or more vehicles, 0 otherwise) *(standard deviation)*	0.436	5.45	1.078 (1.142)	8.23 (10.17)
Threshold (μ)	1.477	25.77	1.582	24.71
Number of observations	2486		2486	
Log-likelihood at zero, LL(0)	−1669.78		−1669.78	
Log-likelihood at convergence, LL(β)	−1504.41		−1498.32	
McFadden pseudo-ρ^2	0.099		0.103	

Table 3
Marginal Effects for the Ordered Probit Model with Random Parameters.

Variable	Marginal Effects of the ordered probit model with random parameters		
	Severe injury	Minor injury	No injury
Month of the year (1 if crash occurred between January and April, 0 otherwise)	−0.0029	−0.0395	0.0424
Median type (1 for raised median, 0 otherwise)	−0.0054	−0.0825	0.0879
Roadway surface condition (1 for dry, 0 otherwise)	0.0045	0.0596	−0.0642
Driver license status (1 for a license from other states or countries, 0 otherwise)	−0.0034	−0.0443	0.0477
Vehicle maneuver just before impending crash (1 if going straight, 0 otherwise)	0.0098	0.1692	−0.1789
Participant level action (1 for lost control of a vehicle, 0 otherwise)	0.0051	0.0615	−0.0666
Participant level safety equipment use (1 if seatbelt was fastened, 0 otherwise)	−0.0066	−0.0736	0.0803
Crash type (1 for overturn, 0 otherwise)	0.0083	0.0871	−0.0954
Alcohol not a factor (1 for no alcohol involved, 0 otherwise)	−0.0303	−0.2012	0.2315
Speed not a factor (1 for non-speed involved, 0 otherwise)	−0.0097	−0.0935	0.1032
Roadway characteristics (1 for horizontal curve, 0 otherwise)	0.0052	0.0607	−0.0658
Participant level action (1 for fatigue, 0 otherwise)	0.0066	0.0692	−0.0758
Number of vehicles involved in the crash (1 if two or more vehicles, 0 otherwise)	−0.0135	−0.2104	0.2239

fixed parameters. Furthermore, the goodness of fit, which is provided at the bottom of Table 2, proves that the random parameter is statistically superior.

The marginal effects, which are illustrated in Table 3, provide additional insights with regard to what occurs with interior injury severity categories, their corresponding probabilities, and the magnitude of change. Table 3 presents further details on the probability and the sign corresponding to each interior category. With regard to the interpretation of the marginal effects for ROR crashes, such as the indicator variable representing roadway surface condition (1 for dry, 0 otherwise), the probability of no injury outcome on average is −0.0642, which means that the probability of sustaining no injury outcome where ROR crashes occur on dry roadway surfaces decreases by 0.0642, on opposed to the probabilities of severe- and minor-injury outcomes that have a positive marginal effect values as illustrated in Table 3.

With regard to the results found in Table 2, environment, roadway, human (person), vehicle, and crash-related factors highlight key findings from the estimated ordered random parameter probit model. All estimated parameters included in the model were found to be statistically significant, with plausible signs. Five parameters were found to be random, with statistically significant standard deviations for the assumed distribution, which is the normal distribution. These variables were crashes that occurred between January and April, raised median type, loss of control of a vehicle, the indicator variable of speed not involved, and the indicator variable of two vehicles or more involved in the crashes.

4.1. Environmental and roadway-related factors

As shown in Table 2, the indicator variable of crashes that occurred between January and April was found to be statistically significant and random, with a mean of 0.161 and a standard deviation of 0.268. This finding indicates that 27.4% of ROR crashes involving a large truck are less than zero when these crashes occurred between January and April, whereas 72.6% of these crashes are greater than zero. That is, 27.4% of ROR crashes that involved large trucks are less likely to result in no injury outcome, whereas 72.6% of these crashes increase the probability of no injury outcome. As seen from the marginal effects on Table 3, the probability of no injury outcome on average is 0.0424, which is greater than the probabilities of severe- and minor-injury outcomes. This finding may be due to the inclement weather, which is more likely to occur between the months of January and April in Oregon, causing more cautious driving. The finding is consistent with previous research performed by Maze et al. (2005) and McLaughlin et al. (2009), where it was demonstrated that environmental conditions increase the probability of ROR crashes involving large trucks.

The indicator variable representing speed not involved was also found to be random, with a mean of 0.335 and standard deviation of 0.133. These values suggest that 0.6% of ROR crashes where speed was not involved are less than zero, whereas 99.4% of these crashes are greater than zero. In other words, 0.6% of ROR crashes involving large trucks where speed was not involved are less likely to result in no injury outcome, whereas 99.4% of these crashes are more likely to end up with no injury outcome. Turning to Table 3, the probability of no injury outcome on average is 0.1032, while the probabilities of severe- and minor-injury outcomes are less than the probability of no injury.

The raised median is a traffic-calming device that is implemented to reduce vehicles' speed. However, in some cases, it correlates with high crash severity, particularly for errant vehicles. In this study, the indicator variable representing raised median was found to be a random parameter with a mean of 0.361 and standard deviation of 0.448. These values indicate that 21.0% of ROR crashes are less than zero when these crashes occurred on a roadway with raised median, whereas 79.0% of ROR crashes under the same conditions are greater than zero. Conversely, 21.0% of ROR crashes involving large trucks that took place on roadways with raised medians are less likely to result in no injury outcome, whereas 79.0% of these crashes increase the probability of no injury outcome. Table 3 shows that the probability of no injury outcome is 0.0879, which is high compared to the probabilities of severe- and minor-injury outcomes. Previous research also found that the presence of a raised median was more likely to decrease crash severity. For instance, Schultz et al. (2011) stated that installing such medians on Utah roadways reduced severe crashes by 36%. Likewise, Alluri et al. (2014) found that converting two-way left-turn lanes to raised medians on Florida's roadways was associated with a 30% reduction in crash rates. However, the randomness of this variable suggests that, for some observations, raised medians may lead to an increased probability of severe injuries.

Horizontal curves are one of the most significant contributory factors to ROR crashes. The possibility of departing the roadway at a curved section is higher than for a straight roadway section. As seen from Table 2, the indicator for the horizontal curve was found to be a fixed parameter, indicating that ROR crashes occurring at horizontal curves were more likely to lead to injuries that were more serious. Table 3 illustrates that no injury outcome will be decreased by −0.0658 when ROR crashes involving large trucks occurs on horizontal curves. Torbic et al. (2004) state that crashes on curved roadway sections are three times more frequent than those occurring on straight roadway sections. They also found that approximately 76% of fatal crashes on curved roadway sections were single ROR crashes. Islam and Hernandez (2013b) found similar results for large-truck crashes.

4.2. Human-related factors

Turning to human-related factors, the indicator variable representing driver license status (1 for a license from other states or countries, 0 otherwise) was found to be statistically significant and a fixed parameter. The marginal effects in Table 3 indicate that for large truck drivers from other states, the probability of experiencing no injury outcome is 0.0477, which is higher than the probabilities of severe and minor injury outcomes. The possible reasons may be due to driver familiarity with Oregon highways, and this parameter estimate may also be capturing the driving complexities related to the diverse geographical nature of the state of Oregon.

The use of safety equipment (1 if a seatbelt was fastened, 0 otherwise) was also found to be significant. This fixed parameter suggests that the probability of no injury outcome on average is 0.0803 higher (see Table 3), while the probabilities of severe injury and minor injury are lower. Islam and Hernandez (2013b) also found the use of seatbelts to be a fixed parameter and that it increased the probability of experiencing no injury.

The parameter estimate representing being fatigued before the crash or not was found to be statistically significant and a fixed parameter. As illustrated in Table 3, the probability of no injury outcome on average is reduced by −0.0758 when ROR crashes involving large trucks occurred due to fatigued drivers. This finding is consistent with (Peng and Boyle, 2012). In their study, they concluded that drowsiness and fatigue were associated with severe and fatal ROR crashes.

The indicator variable for no alcohol involved (i.e., alcohol not a factor) before the crash was found to be significant for ROR crashes and a fixed parameter. Turning to Table 3, the probability of no injury outcome on average is 0.2315, which is higher than the probabilities of severe- and minor-injury outcomes.

The lost control of vehicle indicator variable was found to be statistically significant and a random parameter with a mean of −0.239 and standard deviation of 0.141. These values suggest that 4.5% of ROR crashes involving large trucks where the driver loss control of the vehicle are greater than zero, while 95.5% of these crashes are less than zero. In other words, 4.5% of ROR crashes involving large trucks that occurred due to loss of control of a vehicle are more likely to result in no injury outcome, whereas 95.5% of these crashes are less likely to end up with no injury outcome. Turning to Table 3, this suggests that for most ROR crash occurrences—and taking into account the randomness of this parameter—the probability of no injury outcome on average is reduced by −0.0666. This variable may be capturing driver complexities related to vehicle performance issues, such as a flat tire, or capturing unobserved factors related to driver inattentiveness to roadway environment, hence the randomness of this variable.

4.3. Vehicle and crash-related factors

With respect to the influence of vehicle and crash-related factors on the probability of ROR crash occurrence, the following variables were found to statistically significant. The indicator variable of two or more vehicles involved in the ROR crashes was found to be a random parameter with a mean of 1.078 and a standard deviation of 1.142. These values give parameters of less than zero for 17.3% of ROR crashes involving multiple vehicles and greater than zero for 82.7%. Specifically, 17.3% of ROR crashes involving two or more vehicles are less likely to result in no injury outcome, whereas 82.7% of these crashes increase the probability of no injury outcome. As can be seen from Table 3, the probability of sustaining no injury in ROR crashes involving two or more vehicles will be increased by 0.2239, which is higher than the probabilities of severe- and minor-injury outcomes. This finding may be capturing the effect of vehicle body type in reducing the impact of injury sustained by large-truck drivers (Eluru et al., 2010). Again, given the randomness of this parameter estimate, for a small portion of the ROR crash occurrences, the opposite is true.

With regard to the crash type indicator (1 for overturning, 0 otherwise), it was found to be significant and a fixed parameter. Table 3 shows that the probability of no injury outcome on average is −0.0954, which is lower than the probabilities for severe- and minor-injury outcomes.

The indicator variable of driving straight as an evasive maneuver just before an impending crash was found to be statistically significant and fixed across observations. The marginal effects of driving straight as an evasive maneuver just before an impending crash indicate that the probability of no injury outcome on average is reduced by −0.1789 (see Table 3). This finding can be substantiated by the drivers' expectancy that they will not experience a crash on a straight roadway and it may related to driver behavior (e.g., driver inattentiveness, vehicle performance issues).

5. Conclusions and future research

The current study explores possible contributory factors to ROR crashes that involved large trucks in Oregon, utilizing an ordered random parameter probit modeling framework. The ordered random parameter probit model is an important approach because it provides a mechanism to account and correct for unobserved heterogeneity that can arise from factors related to the driver, vehicle, road environment, weather, variations in police reporting, temporal, and other unobserved factors not captured. The data used in this study comprised crash reports taken from the state of Oregon for the years of 2007–2013, to the best of our knowledge a first with respect to explicitly modeling large-truck-injury severity of ROR crashes.

The results of the analyses performed provided some interesting findings. First, five parameter estimates were found to be random and varied across the ROR crash observations. These were factors related to crashes that occurred between January and April, raised median type, loss of control of a vehicle, the indicator variable of speed not involved, and the indicator variable of two vehicles or more involved in the crashes. However, driver license status, seatbelt use, crash type, alcohol not a factor, the presence of a horizontal curve, and driver fatigue were found to be fixed parameters for the ROR crashes. Looking more closely at the results of some of the significant variables, large-truck drivers who are not licensed in Oregon have a higher probability of experiencing no injury ROR crash outcomes. Another possible explanation for this outcome could be due to driver turnover suffered by the trucking industry and its impact on driver network familiarity. Second, it was discovered that median type (i.e., raised median) increases the probability of no injury outcome in ROR crashes. Third, human-related factors such as fatigued drivers have a higher probability of severe and minor injury. These findings are important from a trucking perspective because these contributory factors can be targeted through firm intervention and continued training.

Although the research performed is exploratory in nature, the ordered random parameter probit modeling framework presented in this work offers a flexible and practical methodology to analyze ROR crashes involving large trucks and to account for unobserved heterogeneity. Furthermore, this study provides information that can aid safety planners and the trucking industry in identifying appropriate countermeasures to help mitigate the number and severity of large-truck ROR crashes. Using the same approach and comparing it with recent statistical models, with an expanded sample of ROR large-truck crashes could provide important new insights into large-truck driving behavior. For example, datasets with driver skill and other cognitive processing information, car-following dynamics, and human response can greatly improve parameter estimates as well as help improve truck-driver training for collision avoidance.

Acknowledgments

We would like to acknowledge the Oregon Department of Transportation, specifically, the Transportation Data Department, for providing the crash dataset and for helping in our understanding of the data. And we would like to also acknowledge Jason Anderson for his help.

References

Abdel-Aty, M., Keller, J., 2005. Exploring the overall and specific crash severity levels at signalized intersections. Accid. Anal. Prev. 37, 417–425.

Abdel-Aty, M., 2003. Analysis of driver injury severity levels at multiple locations using ordered probit models. J. Safety Res. 34, 597–603.

Agbelie, B.R.D.K., 2014. An empirical analysis of three econometric frameworks for evaluating economic impacts of transportation infrastructure expenditures across countries. Transp. Policy 35, 304–310.

Alluri, P., Gan, A., Haleem, K., 2014. Safety impacts of converting two-way left-turn lanes to raised medians and associated design concerns. Transportation Research Board 93rd Annual Meeting.

Anarkooli, A.J., Hosseinlou, M.H., 2016. Analysis of the injury severity of crashes by considering different lighting conditions on two-lane rural roads. J. Safety Res. 56, 57–65.

Anastasopoulos, P.C., Mannering, F.L., 2009. A note on modeling vehicle accident frequencies with random-parameters count models. Accid. Anal. Prev. 41, 153–159.

Aram, A., 2010. Effective safety factors on horizontal curves of two-lane highways. J. Appl. Sci. 10, 2814–2822.

Bhat, C.R., 2003. Simulation estimation of mixed discrete choice models using randomized and scrambled Halton sequences. Transp. Res. Part B: Methodol. 37, 837–855.

Blincoe, L., Miller, T.R., Zaloshnja, E., Lawrence, B.A., 2015. The Economic and Societal Impact Of Motor Vehicle Crashes, 2010 (Revised). National Highway Traffic Safety Administration, Washington, DC.

Cerwick, D.M., Gkritza, K., Shaheed, M.S., Hans, Z., 2014. A comparison of the mixed logit and latent class methods for crash severity analysis. Anal. Methods Accid. Res. 3–4, 11–27.

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.

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.

Compton, R., Berning, A., 2009. Results of the 2007 National Roadside Survey of Alcohol and Drug Use by Drivers. National Highway Traffic Safety Administration, Washington, DC.

Davis, G., Davuluri, S., Pei, J., 2006. Speed as a risk factor in serious run-off-road crashes: bayesian case-control analysis with case speed uncertainty. J. Transp. Stat. 9, 17–28.

Dissanayake, S., 2003. Young drivers and run-off-the-road crashes. In: In: Proceedings of the 2003 Mid-Continent Transportation Research Symposium, Ames, Iowa.

Duncan, C.S., Khattak, A.J., Council, F.M., 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.

Eluru, N., Yasmin, S., 2015. A note on generalized ordered outcome models. Anal. Methods Accid. Res. 8, 1–6.

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.

Eluru, N., Paleti, R., Pendyala, R.M., Bhat, C.R., 2010. Modeling injury severity of multiple occupants of vehicles copula-based multivariate approach. Transp. Res. Rec., 1–11.

Federal Motor Carrier Safety Administration, 2016. Large Truck and Bus Crash Facts 2014. Federal Motor Carrier Safety Administration.

Greene, W.H., 2007. Econometric Analysis, 5th ed. Prentice Hall, Upper Saddle River, NJ.

Haleem, K., Abdel-Aty, M., 2010. Examining traffic crash injury severity at unsignalized intersections. J. Safety Res. 41, 347–357.

Haleem, K., Gan, A., 2013. Effect of driver' s age and side of impact on crash severity along urban freeways: a mixed logit approach. J. Safety Res. 46, 67–76.

Hensher, D.A., Rose, J.M., Greene, W.H., 2015. Applied Choice Analysis, 2nd ed. Cambridge University Press, Cambridge CB2 8BS, United Kingdom.

Islam, M., Hernandez, S., 2013a. Modeling injury outcomes of crashes involving heavy vehicles on texas highways. Transp. Res. Rec.: J. Transp. Res. Board 2388, 28–36.

Islam, M., Hernandez, S., 2013b. Large truck–involved crashes: exploratory injury severity analysis. J. Transp. Eng. 139, 596–604.

Islam, M., Hernandez, S., 2015. Fatality rates for crashes involving heavy vehicles on highways: a random parameter tobit regression approach. J. Transp. Saf. Secur. 8, 247–265.

Islam, S., Jones, S.L., Dye, D., 2014. Comprehensive analysis of single- and multi-vehicle large truck at-fault crashes on rural and urban roadways in Alabama. Accid. Anal. Prev. 67, 148–158.

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.

LeRoy, A.A., Pharm, D., Morse, M.L., 2008. Multiple medications and vehicle crashes: analysis of databases. Traffic Saf. Facts\rTraffic Tech–Technol. Transf. Ser.

Lee, J., Mannering, F., 2002. Impact of roadside features on the frequency and severity of run-off-roadway accidents: an empirical analysis. Accid. Anal. Prev. 34, 149–161.

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.

Liu, C., Subramanian, R., 2009. Factors Related to Fatal Single-Vehicle Run-Off-Road Crashes. Mathematical Analysis Division, National Center for Statistics and Analysis National Highway Traffic Safety Administration.

Mannering, F.L., Bhat, C.R., 2014. Analytic methods in accident research: methodological frontier and future directions. Anal. Methods Accid. Res. 1, 1–22.

Maze, T.H., Agarwal, M., Burchett, G., 2005. Whether Weather Matters to Traffic Demand, Traffic Safety, and Traffic Flow Aurora Program. Center for Transportation Research and Education, Iowa State University.

McGinnis, R.G., Davis, M.J., Hathaway, E.A., 2001. Longitudinal analysis of fatal run-off-road crashes, 1975 to 1997. Highway Saf.: Model. Anal. Manage. Stat. Methods Crash Location, 47–58.

McLaughlin, S.B., Hankey, J.M., Klauer, S.G., Dingus, T.A., 2009. Contributing Factors to Run-Off-Road Crashes and Near-Crashes.

National Highway Traffic Safety Administration (NHTSA), 2012. Traffic Safety Facts: 2011 Data: Alcohol-Impaired Driving. National Highway Traffic Safety Administration (NHTSA).
Neuman, T., Pfefer, R., Slack, K.L., McGee, H., Prothe, L., Eccles, K., Council, F., 2003. NCHRP REPORT 500: Guidance for Implementation of the AASHTO Strategic Highway Safety Plan.Volume 4: A Guide for Addressing Head-On Collisions. Washington, DC.
Obeng, K., 2011. Gender differences in injury severity risks in crashes at signalized intersections. Accid. Anal. Prev. 43, 1521–1531.
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.
Peng, Y., Boyle, L.N., 2012. Commercial driver factors in run-off-road crashes. Transp. Res. Rec.: J. Transp. Res. Board 2281, 128–132.
Quddus, M.A., Noland, R.B., Chin, H.C., 2002. An analysis of motorcycle injury and vehicle damage severity using ordered probit models. J. Safety Res. 33, 445–462.
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, 4013016.
Roy, U., Dissanayake, S., 2011. Comparison of factors associated with run-off-road and non-run-off road crashes in Kansas. J. Transp. Res. Forum 50, 69–86.
Russo, B.J., Savolainen, P.T., Schneider IV, W.H., Anastasopoulos, P.C., 2014. Comparison of factors affecting injury severity in angle collisions by fault status using a random parameters bivariate ordered probit model. Anal. Methods Accid. Res. 2, 21–29.
Savolainen, P.T., Mannering, F.L., Lord, D., Quddus, M.A., 2011. The statistical analysis of highway crash-injury severities: a review and assessment of methodological alternatives. Accid. Anal. Prev. 43, 1666–1676.
Schultz, G.G., Thurgood, D.J., Olsen, A.N., Reese, C.S., 2011. Analyzing raised median safety impacts using bayesian methods. Transp. Res. Rec. 2223, 96–103.
Srinivasan, K.K., 2002. Injury severity analysis with variable and correlated thresholds: ordered mixed logit formulation. Transp. Res. Rec.: J. Transp. Res. Board 1784, 132–141.
Torbic, D.J., Harwood, D.W., Gilmore, D.K., Pfefer, R., Neuman, T.R., Slack, K.L., Hardy, K.K., 2004. NCHRP REPORT 500: Guidance for Implementation of the AASHTO Strategic Highway Safety Plan. Volume 7: A Guide for Reducing Collisions on Horizontal Curves. Washington, DC.
Train, K.E., 1998. Recreation demand models with taste differences over people. Land Econ. 74, 1–23.
Washington, S.P., Karlaftis, M.G., Mannering, F.L., 2011. Statistical and Econometric Methods for Transportation Data Analysis, 2nd ed. Taylor and Francis, Chapman & Hall/CRC, Boca Raton, FL.
Yasmin, S., Eluru, N., 2013. Evaluating alternate discrete outcome frameworks for modeling crash injury severity. Accid. Anal. Prev. 59, 506–521.