Accident Analysis and Prevention 75 (2015) 155–163

Contents lists available at ScienceDirect

Accident Analysis and Prevention

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

A time of day analysis of crashes involving large trucks in urban areas

Jasmine Pahukula [a], Salvador Hernandez [a,*], Avinash Unnikrishnan [b]

[a] School of Civil and Construction Engineering Oregon State University, 309 Owen Hall, Corvallis, OR 97331-3212, USA
[b] Department of Civil and Environmental Engineering, West Virginia University, Room 621 ESB, P. O. Box 6103, Morgantown, WV 26506-6103, USA

ARTICLE INFO

Article history:
Received 14 October 2014
Received in revised form 26 November 2014
Accepted 26 November 2014
Available online xxx

Keywords:
Truck accidents
Injury severity
Mixed logit
Time-of-day
Interstate
Freight

ABSTRACT

Previous studies have looked at different factors that contribute to large truck-involved crashes, however a detailed analysis considering the specific effects of time of day is lacking. Using the Crash Records Information System (CRIS) database in Texas, large truck-involved crashes occurring on urban freeways between 2006 and 2010 were separated into five time periods (i.e., early morning, morning, mid-day, afternoon and evening). A series of log likelihood ratio tests were conducted to validate that five separate random parameters logit models by time of day were warranted. The outcomes of each time of day model show major differences in both the combination of variables included in each model and the magnitude of impact of those variables. These differences show that the different time periods do in fact have different contributing factors to each injury severity further highlighting the importance of examining crashes based on time of day. Traffic flow, light conditions, surface conditions, time of year and percentage of trucks on the road were found as key differences between the time periods.

1. Introduction

1.1. Motivation

Large truck crashes have a considerable impact on society and the economy. It has been estimated that the average cost of no injury crashes (i.e., property damage only), non-fatal injury crashes, and fatal crashes involving large trucks are $15,114, $195,258 and $3,604,518, respectively (Zaloshnja and Miller, 2006). These estimates include medical costs, emergency services costs, property damage costs, lost productivity and monetized value of the pain, suffering and quality of life lost due to death or injury. The estimated cost of large truck crashes between 1997 and 1999 exceeded US$ 19.6 billion (Zaloshnja and Miller, 2004). From the perspective of moving freight, in 2010 it was estimated that large trucks carried roughly 68% of freight tonnage in the U.S totaling approximately 12,500 millions of tons (Federal Highway Administration, 2013). The National Highway Traffic Safety Administration (NHTSA) reports that tonnage is expected to increase by 1.4% per year till 2040 (Federal Highway Administration, 2013). Currently this tonnage is being moved continuously day and night and as the tonnage grows so will the daily distribution of the freight movements required to haul this extra tonnage. This has raised concerns especially in large populated urban areas where congestion is only getting worse, where large truck crashes at various times of the day have created havoc to commutes. The added congestion to these urban commutes is the equivalent of 1.9% of the $14.96 trillion U.S. gross domestic product (GDP) in 2010 (Kilcarr, 2014). Evidently, efforts to improve our understanding of the factors that influence large truck-involved crashes are needed especially from a time of day perspective.

Although there have been several efforts to understand large truck-involved crashes, the relationships between crash related factors, crash severity and time of day effects are still not completely understood. A reason for this stems from the availability of sufficient data to capture the complex interactions of multiple factors under a single framework for various times of day scenarios. Recent studies conducted by (Islam and Hernandez, 2013a,b,b) developed random parameters models to predict injury severity of large truck-involved crashes with data from the Texas Crash Records Information System (CRIS), but considered time of day as a contributing factor. To better understand the relationships of crash related factors and crash severity by time of day separately, the CRIS database is utilized for this study.

In order to clearly identify injury related large truck crash factors, the data set will be divided by land use (i.e., rural and urban) and then further divided into time periods. Khorashadi et al. (2005) identified significant differences between urban and rural crashes due to differing driver, vehicle, environmental, road geometry and traffic characteristics. Additionally, time of day

* Corresponding author. Tel.: +1 541 737 1757.
E-mail addresses: pahukulj@onid.oregonstate.edu (J. Pahukula), sal.hernandez@oregonstate.edu (S. Hernandez), avinashu@gmail.com (A. Unnikrishnan).

http://dx.doi.org/10.1016/j.aap.2014.11.021
0001-4575/

has been identified as a significant factor in previous studies (Islam and Hernandez, 2013a,b,b). Past works capture the impact of time of day by using indicator variables representing various times of day as independent variables in regression models. However there is a complex interaction between variables in these types of models. For example, traffic patterns, light conditions and driver behavior can vary throughout the day. The impact of traffic levels in urban areas during morning time period and afternoon time period on truck injury severity may potentially be different. With this in mind, this study aims to analyze injury crash severity of large truck-involved crashes under an urban land use context and varying time of day scenarios through an econometric modeling approach by developing separate models for five time of days – early morning, morning, mid-day, afternoon and evening. Separate models for different time of day can help pinpoint specific issues.

The random parameters logit (or mixed logit) model is utilize here to gain a better understanding of the complex interactions between those factors found in the dataset and those unobserved factors that may be influencing (i.e., through unobserved heterogeneity). A latent class approach can also account for possible unobserved heterogeneity without having to make an assumption about the parameter distribution which may not always be consistent across all observations. Latent class models can account for possible unobserved heterogeneity by assuming that observations come from distinct classes based on common characteristics. However, one drawback of this approach is the number of classes is usually quite small so there is a coarse approximation of the distribution of heterogeneity (Behnood, 2014). Xiong and Mannering (2013) and Shaheed and Gkritza (2014) have identified another drawback that latent class models do not account for potential variation within a class. Xiong and Mannering (2013) further point out the difficulty in determining the statistically superior model which can vary by dataset. The random parameters approach will be utilized to this dataset to account for the unobserved heterogeneity. To the best of the authors' knowledge, this is the first attempt at modeling injury severity for large truck-involved crashes using a random parameters logit approach on urban freeways by separating crashes by time of day on three injury severity levels (serious injury, minor injury and no injury).

The remainder of the paper is organized as follows. First, a review of the current literature is presented followed by a discussion of the empirical settings and descriptive statistics. Next, the methodological approach is explained and the results are summarized. Finally, implications of the findings and the conclusion are presented.

1.2. Background

Although not the focus of this study, the following references provide valuable insights on time-of-day and its relation to crash rates and injuries sustained during crashes involving large trucks. According to the Fatality Facts provided by Insurance Institute for Highway Safety, highest incidence of deaths due to large truck crashes, nearly 19%, occur between the noon to 3 p.m. period (Fatality Facts 2004: Large Trucks, 2004). Blower and Campbell (1998) analyzed the Fatality Analysis Reporting System (FARS) data set from 1993 to 1995 and found that the higher fatalities occurred during daylight hours. However when fatality rates were calculated, a higher probability of fatality given the occurrence of a crash was observed during night time. An analysis of the General Estimates System (GES) data set for the same period revealed that while there were fewer crashes between midnight and 7 a.m., the chances of severe injuries were higher if a crash occurred during that period. It is important to note here that not all transportation facilities experience the same amounts of vehicular flows, thus exposure to higher traffic volumes may produce varying results with regards to maximum injury severity potential. Other possible exposure variables such as night-time hours of driving, truck-miles traveled, or ton-miles when considered could provide additional information on severity rates of large-truck involved crashes. In future work, the authors are examining methods that take into account exposure based data and crash analysis techniques for large-truck crashes.

Curnow (2002) analyzed the Australian Truck Crash Database and found that articulated truck crash incidents were spread evenly throughout the 24 h period whereas majority of the rigid truck crashes occurred during the day. Ghariani (2001) studied ten years of truck crash data from 1991 to 1999 obtained from Texas Department of Public Safety and found that a significant majority of the crashes occurred during day time. Similar trends were found in the rural freeways of Wyoming and Nebraska for the year 2000–2009 (Offei and Young, 2014). Knipling and Bocanegra (2008) analyzed the frequency of crash occurrence of combination unit trucks and single unit trucks from the truck crash causation study data (LTCCS) and found that the majority of the crashes occurred during the day and especially during rush hours. The percentage of crashes was found to be higher under dark conditions for combination unit trucks compared to single unit trucks. A majority of the above insights which focus on frequencies and distribution of crash occurrence based on time of day can be explained by the fact that most truck operations occur during the day.

Duncan et al. (1998) used an ordered probit model to understand the factors affecting truck–car rear end collisions based on highway safety information system data in North Carolina from 1993 to 1995. Injury severities were found to be higher during night time. Chang and Mannering (1999) analyzed the accidents in King County using a Nested Logit Model and found that for truck involved accidents there is a 50% higher chance of an injury or fatality if the accident occurred during night time and a 37% decrease in the probability of a possible injury if the accident occurred during night time.

Khorashadi et al. (2005) used a multinomial logit structure to understand the differences in factors affecting the severities of large-truck involved accidents in urban and rural areas using four years of crash data from 1997 to 2000 maintained by California Department of Transportation. The multinomial logit specifications were preferred to several nested logit specifications. Darker driving conditions were found to increase the probability of severe or fatal injury crashes. The probability of severe or fatal injury crashes decreased during rush hour with the decrease more prominent in the morning rush hour. Zhu and Srinivasan (2011) used an ordered probit model on the LTCCS data and found that crashes which occurred between 7:30 p.m. to 6:00 a.m lead to more severe crashes.

Lemp et al. (2011) used the heteroskedastic ordered probit model on the LTCCS dataset to study the impact of vehicle, environmental, and crash level variables on vehicle based and crash based maximum injury severity and found that non-bright conditions increased the probability of fatality. Chen and Chen (2011) studied the impact of driver, vehicle, environmental, roadway, temporal, and accident characteristics on single vehicle and multiple vehicle accidents involving large trucks using the highway safety information system data set for the state of Illinois from 1991 to 2000. Mixed logit specification was found to be better than the multinomial logit model. The probability of possible injury/non-incapacitating injury was found to increase during rush hour in single vehicle model. Non-bright conditions were found to significantly increase the probability of injury or fatality in multi-vehicle accident case.

Islam and Hernandez (2013b) used a random parameter ordered probit specification to study the impact of human, vehicle, and road environmental factors on large truck crash injury severity using the National Automotive Sampling System General Estimated System (NASS-GES) database from 2005 to 2008. In contrast to the insight obtained from other literature, the likelihood of lower injury severity was higher when crashes occurred in darker conditions. In another research effort by Islam and Hernandez (2013a), they developed mixed logit models for the truck crashes in Texas using data from the Texas Peace Officer's Crash Reports database for the year 2006–2010. The likelihood of fatal, incapacitating and possible injuries was found to reduce during the afternoon peak period due to congestion effects. The likelihood of fatal and incapacitating injuries increased during dark conditions.

The time of day dimension has been studied in automobile crashes using multivariate logistic regression (Martensen and Dupont, 2013), binary regression embedded in a hierarchical Bayesian framework (Qin et al., 2006), binomial regression and classification and regression tree based analysis (Chang and Chen, 2005) Almost all the models developed above use indicator variables to study the impact of time of day or lighting conditions on crash injury severity. However, such an approach is limited as different variables interact with each other and affect the injury severity outcomes in a complex and different manner depending on time of day. For example, driver behavior will be significantly different in the morning peak compared to the afternoon peak. The traffic flow obviously varies during the peak and off peak periods. A simple indicator variable based approach will not properly account for the complexities of the interactions during the different time periods. In order to account for these changes, it is critical to develop separate models so that the accurate impact of driver, environmental, and roadway related factors on injury severities and their variations with time of day can be estimated accurately. This paper adopts the methodology of Morgan and Mannering (2011) by estimating separate models for time of day.

2. Method

2.1. Data

Large truck crashes between 2006 and 2010 reported by Texas Peace Officer's Crash Reports were utilized in this study. Only large truck-involved crashes on urban roadways were considered. A sample of 11,560 data observations were extracted from the CRIS database. Each observation represents the maximum level of injury sustained by the driver. Three different data components (crash, vehicle and person) were linked based on the 'Crash ID'.

Due to low data observations for the higher injury severity outcomes the five injury severity outcomes as defined by the KABCO injury scale were grouped into three categories (severe injury, minor injury and no injury). Serious injuries included fatalities and incapacitating injuries while minor injury included non-incapacitating injury and possible injury and property damage only crashes make up the no injury category. Overall, no injury crashes, minor injury crashes and serious injury crashes accounted for 90.8% (N = 10,499), 7.6% (N = 878) and 1.6% (N = 183), respectively. The individual data sets separated by time of day followed the same pattern where no injury crashes had the most observations and serious injury crashes accounted for the lowest percentage of crashes.

The effect of time of day on injury severity is the focus of this study. The analysis examined five different time periods, as shown in Table 1, which shows descriptive statistics of key variables included in the five models.

The driver demographics including gender, age and restraint use remain consistent throughout the different time periods. Only 13.6% of the crashes occurred during dark lighting conditions. Male drivers accounted for about 93% of the total observations for each of the five datasets. Drivers under the age of 25 and between 35 and 45 accounted for about 10% and 29% of the total observations, respectively. Drivers using both a lap and shoulder belt grossed about 90% of the total observations.

Table 1
Descriptive statistics of key variables by time of day.

Meaning of variable	Early morning (12:00–4:00 a.m.)		Morning (5:00–9:00 a.m.)		Mid-day (10:00 a.m.–3:00 p.m.)		Afternoon (4:00–8:00 p.m.)		Evening (9:00–11:00 p.m.)	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Age (1 if age <25, 0 otherwise)	0.12	0.32	0.09	0.29	0.09	0.29	0.09	0.28	0.09	0.29
Age (1 if age 35–45, 0 otherwise)	–	–	–	–	0.29	0.45	0.30	0.46	–	–
Base type (1 if granular base: flex or stabilized earth, 0 otherwise)	–	–	–	–	0.56	0.50	–	–	0.60	0.90
Collision type (1 if going straight and sideswipe, 0 otherwise)	0.37	0.48	0.42	0.49	0.41	0.49	0.44	0.50	0.49	0.50
Contributing factor (1 if failed to control speed, 0 otherwise)	–	–	–	–	0.08	0.27	–	–	–	–
Contributing factor (1 if unsafe to change lanes, 0 otherwise)	–	–	–	–	0.09	0.28	–	–	–	–
Gender (1 if male, 0 otherwise)	0.93	0.26	0.94	0.23	0.94	0.24	0.93	0.26	0.92	0.27
Intersection related (1 if at intersection, 0 otherwise)	–	–	0.28	0.45	0.32	0.47	0.26	0.44	0.24	0.42
Light condition (1 if dark including dawn and dusk, 0 otherwise)	0.93	0.26	–	–	–	–	–	–	–	–
Median width (1 if width between 51 and 75 feet, 0 otherwise)	–	–	–	–	0.14	0.35	–	–	–	–
Month (1 if crash occurred between June and August, 0 otherwise)	–	–	–	–	–	–	0.27	0.44	–	–
Object struck (1 if another vehicle, 0 otherwise)	–	–	–	–	0.86	0.35	0.90	0.31	–	–
Percentage of trucks (1 if percent trucks between 12% and 16%, 0 otherwise)	–	–	0.20	0.40	–	–	–	–	–	–
Percentage of trucks (1 if more than 16% trucks, 0 otherwise)	–	–	–	–	–	–	–	–	0.12	0.33
Restraint use (1 if used shoulder and lap belt, 0 otherwise)	0.87	0.34	0.91	0.28	0.91	0.28	0.91	0.28	0.90	0.30
Right shoulder width (1 if width 20 feet, 0 otherwise)	0.60	0.49	–	–	–	–	–	–	–	–
Right shoulder width (1 if width greater than 20 feet, 0 otherwise)	–	–	–	–	–	–	–	–	0.23	0.42
Road alignment (1 if level and straight, 0 otherwise)	0.73	0.45			0.78	0.42	–	–	–	–
Surface condition (1 if dry at the time of the crash, 0 otherwise)	–	–	–	–	–	–	0.86	0.35	–	–
Vehicle maneuver before the crash (1 if changing lanes, 0 otherwise)	–	–	–	–	0.09	0.28	–	–	–	–
Vehicle maneuver before the crash (1 if going straight, 0 otherwise)	–	–	0.12	0.33	–	–	–	–	–	–
Weather condition (1 if clear at the time of the crash, 0 otherwise)	–	–	–	–	–	–	–	–	0.85	0.36
Weather condition (1 if raining at the time of the crash, 0 otherwise)	–	–	–	–	–	–	0.09	0.29	–	–

Morning, midday and afternoon data set accounted for 42.4%, 40.5%, and 44.1% of the total observations, respectively.

The crash characteristics specifically a sideswipe collision varied across the five time periods. The evening data set had a high of 48.9% while the early morning data set had a low of 37.2% of the total observations resulting in a sideswipe crash. Sideswipe crashes in the morning, midday and afternoon data set accounted for 42.4%, 40.5%, and 44.1% of the total observations, respectively.

2.2. Modeling approach

As previously mentioned in Section 2.1, each large truck-involved crash (observation) used in this study represents the maximum level of injury sustained by the driver. The injury severity levels are serious injury (fatalities and incapacitating injuries), minor injury (non-incapacitating injury and possible injury) and no injury (property damage only). It has been shown in previous studies that the random parameters logit model is an appropriate method of modeling the ordered nature of injury severity data (Gkritza and Mannering, 2008; Morgan and Mannering, 2011; Islam and Hernandez, 2013a; Islam et al., 2014). The advantage of utilizing this approach is that it overcomes the limitations of previous models(e.g., multinomial logit, nested logit, ordered probit, Bayesian ordered, etc.) by allowing the parameter estimates to be random (Savolainen et al., 2011,b; Islam and Hernandez, 2013a,b). In allowing the parameter estimates to vary across observations, in contrast to fixed parameter models, one can account for some of the unobserved heterogeneity (unobserved factors) and avoid the independence of irreverent alternatives (IIA) property violations (Savolainen et al., 2011; Washington et al., 2010). As a result, any heterogeneous effects and correlation in unobserved factors are addressed. Thus, in this study a random parameters logit modeling approach is used to model injury severity for large truck-involved crashes in urban areas for various time of day scenarios.

To start, a linear function is used to model the relationship between the latent continuous variable for injury severity and the explanatory variables as follows: estimate the injury severity (i.e., serious injury, minor injury and no injury) for the large truck-involved crashes:

$$S_{in} = \beta_i X_{in} + \varepsilon_{in} \tag{1}$$

where S_{in} is the latent continuous variable for injury severity i (where $i \in I$ denotes serious injury, minor injury and no injury) of an individual n (driver), X_{in} is the vector of explanatory variables (or the contributing factors to that injury severity), β_i is the vector of estimated parameters for each injury severity, and ε_{in} is the error term to capture the effects of the unobserved characteristics for each time of day model (Washington et al., 2010). Furthermore, if the ε_{in} values in Eq. (1) are assumed to be generalized extreme distributed, McFadden has shown that the following multinomial logit formulation results are as follows (McFadden, 1981):

$$P_n(i) = \frac{\exp[\beta_i X_{in}]}{\sum_{\forall I} \exp(\beta_I X_{In})} \tag{2}$$

where $P_n(i)$ is the probability of an individual n (driver) suffering injury severity i (where $i \in I$ denotes serious injury, minor injury and no injury).

To account for the possibility of unobserved heterogeneity due to under reporting of crashes and to capture the randomness associated to some of factors necessary to understand injury severity sustained by the drivers, Eq. (2) is extended and the following is the resulting random parameters logit model (Mcfadden and Train, 2000; Train, 2003):

$$P_n(i) = \int \frac{\exp[\beta_i X_{in}]}{\sum_{\forall I} \exp(\beta_I X_{In})} f(\beta_i|\varphi) d\beta_i \tag{3}$$

where, $f(\beta|\varphi)$ is the density function of β_i and φ and is the vector of parameters of the density function (mean and variance). Eq. (3) can now account for injury severity outcome specific variations of the effect of the factors X_{in} on large truck-involved crash probabilities for each time of day model developed, with the density function $f(\beta|\varphi)$ used to determine β_i. The random parameters logit probabilities are then a weighted average for different values of β_i across the observations where some elements of the vector β_i could be fixed and some randomly distributed. If the parameters are found to be random, the random parameter logit weights can be determined by the density function $f(\beta|\varphi)$ (Washington et al., 2010).

To estimate the random parameters logit model as illustrated by Eq. (3), maximum likelihood estimation is performed through a simulation based approach to address the computational complexity of computing the outcome probabilities. The chosen simulation approach utilizes Halton draws which have been shown to provide a more efficient distribution of the draws for numerical integration than purely random draws (Halton, 1960; Train, 1999; Bhat, 2003). The marginal effects are computed for the variable included in the models. The marginal effect shows the effect of a one unit change of variable, x, on the injury outcome i. For marginal effects computations the readers are referred to (Washington et al., 2010)

3. Empirical results

Maximum likelihood and simulation-based maximum likelihood methods are utilized to estimate parameter vector β_i for the full urban and urban time of day random parameters logit models. We considered normal, lognormal, triangular, and uniform distributions for the distribution of the random parameters in our analysis. However, the normal distribution was found to be statistically significant. In addition, to estimate the random parameters, 200 Halton draws were used. This number has been empirically shown to produce accurate parameter estimates under the simulation-based maximum likelihood estimation procedure (Bhat, 2003).

Once the models were developed, log likelihood ratio tests were conducted to determine if separate models based on time of day were justified following the procedures found in (Washington et al., 2010). The full urban model was compared to the individual time of day models with two methods. The first test compared the full model against all of the time of day models while the second test compared the models individually. The first log likelihood ratio test for transferability is as follows:

$$\chi^2 = -2\left[\mathrm{LL}_{\mathrm{Full}}(\beta^{\mathrm{Full}}) - \sum_{j=1}^{J} \mathrm{LL}_j(\beta^j)\right] \tag{4}$$

where $\mathrm{LL}_{\mathrm{Full}}(\beta^{\mathrm{Full}})$ is the log likelihood at convergence of the full model (−3386.73), $\mathrm{LL}_j(\beta^j)$ is the log likelihood at convergence of subgroup j (i.e., the set of time of day periods of early morning, morning, midday, afternoon and evening) using the same variables included in the full model, and J is the total number of subgroups ($\sum_{j=1}^{J} \mathrm{LL}_j(\beta^j) = -3321.16$). The χ^2 statistic ($\chi^2 = -131.1431$), with degrees of freedom equal to the summation of the number of estimated parameters in all time of day models minus the number of estimated parameters in the overall model, provides the confidence level at which we can reject the null hypothesis. The null hypothesis states that there is no difference between the model parameters in the full and separate models (i.e., the parameters are the same) (Washington et al., 2010). The Chi square statistics with 60 degrees of freedom resulted in a value greater

Table 2
Summary of transferability test comparing the individual time of day models (Chi-square statistic and degrees of freedom).

j_1	j_2				
	Early morning	Morning	Mid-day	Afternoon	Evening
Early morning	0.00	800.27 (d.f = 9)	621.24 (d.f = 16)	975.42 (d.f = 11)	796.39 (d.f = 10)
Morning	735.83 (d.f = 9)	0.00	924.02 (d.f = 16)	1394.64 (d.f = 11)	1,169.25 (d.f = 10)
Mid-day	1570.69 (d.f = 9)	2,354.19 (d.f = 9)	0.00	2811.63 (d.f = 11)	2,336.08 (d.f = 10)
Afternoon	644.16 (d.f = 9)	1,017.51 (d.f = 9)	865.66 (d.f = 16)	0.00	1,106.83 (d.f = 10)
Evening	180.04 (d.f = 9)	417.24 (d.f = 9)	366.07 (d.f = 16)	557.09 (d.f = 11)	0.00

than then 99.99% confidence limit ($\chi^2 = 99.16$), indicating that the models have statistically significantly different model parameters.

For further validation a second log likelihood test was conducted to test the transferability of coefficients from the full model to each time of day model. The second log likelihood ratio test for transferability is as follows:

$$\chi^2 = -2[\mathrm{LL}_{j_1 j_2}(\beta^{j_1 j_2}) - \mathrm{LL}_{j_1}(\beta^{j_1})] \quad (5)$$

where $\mathrm{LL}_{j_1 j_2}(\beta^{j_1 j_2})$ is the log likelihood at convergence of a model using the converged parameters from the j_2's model (using j_2's data) on time period j_1's data and $\mathrm{LL}_{j_1}(\beta^{j_1})$ is the log likelihood at convergence of the model using time period j_1's data (without constraining the parameters). The χ^2 statistic with degrees of freedom equal to the number of estimated parameters in $\beta^{j_1 j_2}$ provides the probability that the models have different parameters. The second set of log likelihood ratio tests all yield Chi square statistics greater than the 99.99% confidence limit based on specified degrees of freedom, further validating that separate models by time of day is justified. The results of the second transferability test (Eq. (5)) can be found in Table 2 below.

The results of the log likelihood tests provide statistically significant evidence, at 99.99% confidence levels, that separate severity models by time of day should be estimated. The individual time of day model estimation results were statistically significant within a 95% confidence level and are presented in Tables 3–7 .

4. Discussion

Of the 22 variables included in the time of day models, only four variables were consistent in each time period. Restraint use, sideswipe collision, age less than 25 and male drivers were found to affect the injury severity regardless of the time, still the sign and magnitude of the estimated coefficients vary across the time of day models. For example, the restraint use indicator is positive in the early morning model indicating that using a lap and shoulder restraint will increase the likelihood of a minor injury. This parameter was also found to be random and normally distributed with a mean 3.046 and a standard deviation 3.378. This suggests that 37.4% of the observations have a mean less than zero, or that 37.4% of the observations are less likely to be involved in a minor injury crash. The restraint use indicator included in the afternoon model was found to be negative indicating that using a lap and shoulder restraint will decrease the likelihood of a minor injury. One possible explanation for the difference could be the light conditions and traffic patterns. The early morning time period can be characterized by dark lighting conditions and lower traffic

Table 3
Random parameters logit injury severity model for early morning large truck-involved crashes.

Meaning of variable	Coefficient	*t*-statistic	Marginal effects		
			Severe injury	Minor injury	No injury
Severe injury					
Constant (standard error of parameter distribution)	−3.091 (1.826)	−3.28 (2.07)			
Restraint use (1 if used shoulder and lap belt, 0 otherwise)	−2.237	−3.3	−0.033	0.0027	0.030
Minor injury					
Restraint use (1 if used shoulder and lap belt, 0 otherwise) (standard error of parameter distribution)	−3.046 (3.378)	−3.00 (3.29)	0.0003	0.011	−0.011
Collision type (1 if sideswipe, 0 otherwise)	−2.618	−4.09	0.002	−0.024	0.022
Road alignment (1 if level and straight, 0 otherwise)	−1.262	−3.12	0.004	−0.046	0.041
No injury					
Age group (1 if age less than 25, 0 otherwise)	−1.946	−3.98	0.01	0.017	−0.026
Gender (1 if male, 0 otherwise)	1.526	3.24	−0.025	−0.071	0.096
Light condition (1 if dark including dawn and dusk, 0 otherwise)	−1.745	−3.06	0.033	0.086	−0.119
Right shoulder width (1 if width 20 feet, 0 otherwise)	0.747	2.39	−0.008	−0.022	0.03
Model statistics					
Number of observations	866				
Restricted log-likelihood	−951.398				
Log-likelihood at convergence	−390.055				
McFadden pseudo-*R*-squared (ρ^2)	0.590				

Table 4
Random parameters logit injury severity model for morning large truck-involved crashes.

Meaning of variable	Coefficient	t-statistic	Marginal effects		
			Severe injury	Minor injury	No injury
Severe injury					
Constant	−3.456	−5.86			
Minor injury					
Restraint use (1 if used shoulder and lap belt, 0 otherwise) (standard error of parameter distribution)	−3.230 (2.577)	−3.29 (2.48)	0.004	−0.024	0.02
Collision type (1 if sideswipe, 0 otherwise)	−1.311	−2.9	0.001	−0.01	0.009
Gender (1 if male, 0 otherwise) (standard error of parameter distribution)	−1.117 (4.215)	−0.97 (4.05)	0.0001	0.071	−0.071
No injury					
Age group (1 if age less than 25, 0 otherwise) (standard error of parameter distribution)	−1.657 (4.009)	−2.14 (3.45)	0.01	0.008	−0.018
Gender (1 if male, 0 otherwise)	4.171	4.04	−0.015	−0.074	0.089
Intersection related (1 if at or intersection related including driveway access points, 0 otherwise)	0.961	2.19	−0.001	−0.005	0.006
Vehicle maneuver before the crash (1 if going straight and sideswipe, 0 otherwise) (standard error of parameter distribution)	−2.353 (2.819)	−2.37 (2.17)	0.013	0.014	−0.027
Percentage of trucks (1 if percent trucks between 12% and 16%, 0 otherwise)	0.976	2.05	−0.001	−0.004	0.004
Model statistics					
Number of observations	2,659				
Restricted log-likelihood	−2,921.21				
Log-likelihood at convergence	−682.709				
McFadden pseudo-R-squared (ρ^2)	0.766				

volumes whereas typically the afternoon period is lighted with high traffic volumes. Thus, the combination of reduced travel speeds and increased sight distance could explain the decreased likelihood of being involved in a minor injury crash.

Of the 22 variables included in the time of day models, 11 variables were found to be random and normally distributed. These random variables account for unobserved heterogeneity and indicate that the effect of a particular variable is varied across the observations. In other words, a portion of the observations may have an increased probability of a certain injury severity and the other portion of the observations will have a decreased probability of that injury severity due to that variable. For example, in the morning model the male indicator was found to be significant and random with a normal distribution. The mean is 1.117 and the standard deviation of 4.215 specifying that for 57.9% of the observations the mean is below zero. In other words, 57.9% of the observations have a decreased probability of being involved in a minor injury crash while 42.1% of the observations have an increased probability of being involved in a minor injury crash.

The results of each time of day model show major differences in both the combination of variables included in each model and the magnitude of those variables. These differences show that the

Table 5
Random parameters logit injury severity model for mid-day large truck-involved crashes.

Meaning of variable	Coefficient	t-statistic	Marginal effects		
			Severe injury	Minor injury	No injury
Severe injury					
Constant	−2.302	−4.79			
Base type (1 if flex base or stabilized earth, 0 otherwise)	0.804	2.22	0.007	−0.001	−0.006
Vehicle maneuver before the crash (1 if changing lanes, 0 otherwise)	−1.934	−2.62	−0.001	0.0002	0.001
Road alignment (1 if level and straight, 0 otherwise)	−0.942	−2.79	−0.012	−0.04	0.052
Median width including inside shoulder (1 if width between 51 and 75 feet, 0 otherwise)	1.07	2.98	0.003	−0.0004	0.003
Minor injury					
Restraint use (1 if used shoulder and lap belt, 0 otherwise) (standard error of parameter distribution)	−3.388 (3.721)	−4.55 (5.89)	0.001	0.019	−0.021
Road alignment (1 if level and straight, 0 otherwise)	−0.707	−2.75	−0.012	−0.040	0.052
Gender (1 if male, 0 otherwise) (standard error of parameter distribution)	−0.53 (2.377)	−0.8 (4.03)	−0.0002	0.03	−0.03
Age group (1 if age between 35 and 45, 0 otherwise)	0.635	2.22	0.014	0.036	−0.050
Collision type (1 if sideswipe, 0 otherwise)	−1.504	−4.55	0.001	−0.011	0.011
Contributing factor (1 if failed to control speed, 0 otherwise)	1.04	2.29	−0.0002	0.004	−0.003
No injury					
Age group (1 if age less than 25, 0 otherwise) (standard error of parameter distribution)	-2.347 (2.749)	−5.41 (3.57)	0.017	0.01	−0.027
Gender (1 if male, 0 otherwise)	1.849	5.24	−0.014	−0.042	0.056
Intersection related (1 if at or intersection related including driveway access points, 0 otherwise)	1.208	4.38	−0.003	−0.009	0.011
Object struck (1 if another vehicle, 0 otherwise)	0.945	3.8	−0.006	−0.019	0.025
Vehicle maneuver before the crash (1 if changing lanes, 0 otherwise)	1.5	2.52	−0.0003	−0.002	0.002
Model statistics					
Number of observations	4,571				
Restricted log-likelihood	−5,021.767				
Log-likelihood at convergence	−1,324.138				
McFadden pseudo-R-squared (ρ^2)	0.736				

Table 6
Random parameters logit injury severity model for afternoon large truck-involved crashes.

Meaning of variable	Coefficient	t-statistic	Marginal effects		
			Severe injury	Minor injury	No injury
Severe injury					
Age group (1 if age between 35 and 45, 0 otherwise)	−1.853	−2.47	−0.001	0.0004	0.001
Month (1 if crash occurred between June and August, 0 otherwise)	−1.451	−2.32	−0.002	0.001	0.001
Minor injury					
Constant	3.158	8.06			
Restraint use (1 if used shoulder and lap belt, 0 otherwise)	−1.665	−4.9	−0.013	−0.079	0.092
Collision type (1 if sideswipe, 0 otherwise)	−1.115	−3.96	0.0014	−0.011	0.01
Weather condition (1 if raining at the time of the crash, 0 otherwise)	−1.267	−2.92	0.001	−0.005	0.004
No injury					
Age group (1 if age less than 25, 0 otherwise)	−1.771	−5.04	0.003	0.011	−0.014
Gender (1 if male, 0 otherwise) (standard error of parameter distribution) (standard error of parameter distribution)	3.978 (2.858)	5.22 (4.51)	−0.001	−0.011	0.013
Intersection related (1 if at or intersection related including driveway access points, 0 otherwise)	0.739	2.5	−0.001	−0.006	0.006
Object struck (1 if another vehicle, 0 otherwise)	2.38	6.88	−0.013	−0.058	0.07
Surface condition (1 if dry at the time of the crash, 0 otherwise)	0.827	2.58	−0.004	−0.022	0.026
Model statistics					
Number of observations	2,763				
Restricted log-likelihood	−3,035.466				
Log-likelihood at convergence	−669.006				
McFadden pseudo-R-squared (ρ^2)	0.780				

different time periods do in fact have different contributing factors to each injury severity further highlighting the importance of examining crashes based on time of day.

As presented in Table 5, three variables were found to be significant in only the mid-day model: changing lanes, median width between 51 and 75 feet and speeding. Changing lanes and speeding could be capturing the uncongested conditions of the transportation facility. As a reminder, the mid-day dataset includes crashes between 10:00 a.m. to 3:00 p.m. that is in between the typical morning and afternoon traffic peak volume periods. Since there are lower traffic volumes during this time period truck drivers are able to travel at higher speeds and perhaps even change lanes to pass slower moving trucks. A large median can also increase a driver's comfort level which could result in increased speed.

Variables found to be exclusive to the afternoon model (4:00–8:00 pm) consist of crashes occurring between June and August, during raining conditions and on a dry surface. Crashes occurring during the summer between June and August were found to decrease the likelihood of a severe injury which may be explained by the types of trips made during this time period. Normally this time frame would be considered the 'after school' period, but during the summer there may be fewer students (i.e., young drivers) on the road.

Weather also had an increased impact during the afternoon time frame. Both rain and dry conditions were found to affect the injury severity. Rain at the time of the crash lead to a decreased likelihood of a minor injury crash while a dry surface at the time of the crash lead to an increased likelihood of a no injury crash. One factor that could be influencing injury outcomes that is not

Table 7
Random parameters logit injury severity model for evening large truck-involved crashes.

Meaning of variable	Coefficient	t-statistic	Marginal effects		
			Severe injury	Minor injury	No injury
Severe injury					
Constant	−3.482	−4.47			
Base type (1 if flex base or stabilized earth, 0 otherwise)	−1.612	−2.16	−0.008	0.001	0.007
Gender (1 if male, 0 otherwise)	-1.806	−2.54	−0.016	0.001	0.015
Right shoulder width (1 if width greater than 20 feet, 0 otherwise)	2.083	2.77	0.013	−0.002	−0.011
Percentage of trucks (1 if more than 16% trucks, 0 otherwise)	1.864	2.18	0.006	−0.0004	−0.005
Minor injury					
Restraint use (1 if used shoulder and lap belt, 0 otherwise) (standard error of parameter distribution)	6.152 (5.242)	2.33 (2.61)	0.041	−0.450	0.408
Collision type (1 if sideswipe, 0 otherwise)	−2.276	−3.31	0.001	−0.024	0.023
Weather condition (1 if clear at the time of the crash, 0 otherwise)	−1.044	−2.63	0.002	−0.029	0.028
No injury					
Age group (1 if age less than 25, 0 otherwise)	−2.874	−5.11	0.016	0.017	−0.033
Intersection related (1 if at or intersection related including driveway access points, 0 otherwise)	1.578	2.77	−0.002	−0.011	0.013
Model statistics					
Number of observations	701				
Restricted log-likelihood	−770.127				
Log-likelihood at convergence	−223.83				
McFadden pseudo-R-squared (ρ^2)	0.709				

included in the model is light condition. This time frame is tricky because depending on the season the light condition can change dramatically between 4:00 pm and 8:00 pm. The elasticity estimates suggest that rain at the time of the crash increases the likelihood of a serious injury or no injury. One possible explanation could be the combination of the light condition as well as the weather conditions. For example, rain and dark lighting conditions can significantly reduce the sight distance as well as the friction between the tire and the roadway surface thus increased the possibility of a crash and a potentially higher injury severity.

Clear weather conditions, a large right shoulder (greater than 20 feet), and a high percentage of large trucks (greater than 16%) were the variables unique to the evening model (9:00–11:00 pm). The evening time period is characterized by dark lighting conditions resulting in lower sight distance, usually lower traffic volumes potentially yielding lower speeds, and possibly sleepy drivers that may lead to inattentive driving. The model results reveal that clear weather conditions increase the likelihood of a severe or no injury and decrease the likelihood of a minor injury crash. One possible explanation could be that clear weather conditions maximized the sight distance available during the evening providing drivers with more time to react to an incidence, hence no injury crashes; however increased sight distance can also provide additional confidence for the driver and promote higher speeds leading to severe injury crashes.

A wide shoulder width was found to increase the likelihood of a severe injury. Excessively wide shoulders may promote improper use of the additional space. Instead of providing positive separation from roadside obstacles, vehicles including large trucks could use that space to pull off the roadway briefly to complete a task such as reading a map. Vehicles located in the right shoulder could create obstacles and distractions for drivers.

A high percentage of large trucks on the roadway were found to increase the likelihood of a crash in both the morning and evening period. The morning model found that if the percentage of large trucks falls between 12% and 16% the likelihood of a no injury crash is increased. Driver fatigue due to inadequate sleep while on the road could be a contributing factor to the crash occurrence while congested morning peak conditions could explain the lower injury severity. The evening model found that if the percentage of large trucks exceeds 16% the likelihood of a serious injury increases. Driver fatigue along with reduced sight distance and lower traffic volumes (i.e., increased speed) could contribute to the high injury severity. McCartt et al. (2000) surveyed 593 long-distance truck drivers randomly on the road at select truck stops. The questions were designed to address typical predictors of driver fatigue. The results show that grueling schedules and poor sleep on road were some factors causing long-distance truck drivers to fall asleep on the road. The survey also revealed that some truck drivers exceeded the 10 consecutive driving hour limit and falsified their log books in order to make the delivery on time.

In summary, the results provide insights related to the impact of crash factors and the complex interactions of these factors on crash severity by time of day in urban areas. Additionally, various factors were found to be random and accounting for the presence of unobserved heterogeneity validating the methodological approach of the random parameters logit model.

5. Conclusions

Random parameters logit models are utilized to examine the effect of time of day on the injury severity of large truck-involved crashes. Using crashes on urban freeways between 2006 and 2010 in Texas, it was determined that separate random parameters logit models are warranted. There were three injury severity outcomes: serious injury (fatality and incapacitating injury), minor injury (non-incapacitating and possible injury) and no injury (property damage only) and there were five time periods: early morning (12:00–4:00 a.m.), morning (5:00–9:00 a.m.), mid-day (10:00 a.m.–3:00 p.m.), afternoon (4:00–8:00 p.m.) and evening (9:00–11:00 p.m.).

The results of the individual models demonstrate considerable differences among the five time periods. Key differences include traffic flow, light conditions, surface conditions, time of year, and percentage of trucks on the road. Ever-changing traffic flow patterns throughout the day were evident in the mid-day model. Free-flow like characteristics such as speeding and changing lanes contributed to large truck-involved crashes between 10:00 a.m. and 3:00 p.m. (i.e., typically uncongested time period). The summer indicator variable in the afternoon model may also suggest that traffic volume may impact injury severity. Crashes between June and August were found to decrease the likelihood of a severe injury which may be explained by less young drivers on the road due to summer vacation.

The evening model suggests that clear weather conditions, but dark light conditions results in either a serious injury or no injury crash. The clear weather conditions may promote speeding for some drivers, while other drivers may take more precaution under dark lighting conditions. Finally, a high percentage of large trucks on the roadway increased the likelihood of a crash in the morning and evening. This could be explained by lack of sleep either from the previous night or failing to pull into a rest stop when the driver is fatigued.

Although the results of this study are exploratory, the results themselves provide evidence of the effect of time of day on large truck-involved crashes. In future work, the authors are working on utilizing these results to develop planning tools to help mitigate the impact of these types of crashes. In addition, we are addressing the spatial transferability of the models to other state specific datasets.

References

Behnood, A., 2014. Latent Class Analysis of the Effects of Age Gender, and Alcohol Consumption on Driver-Injury Severities. Forthcoming in Analytic Methods for Accident Research, Revised, October 2, 2014.

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

Blower, D.F., Campbell, K., 1998. Fatalities and injuries in truck crashes by time of day. Report: UMTRI-98-48, Ann Arbor, MI (accessed from: http://deepblue.lib.umich.edu/bitstream/handle/2027.42/1261/91796.0001.001.pdf?sequence=2).

Chang, L.-Y., Chen, W.-C., 2005. Data mining of tree-based models to analyze freeway accident frequency. J. Safety Res. 36, 365–375.

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.

Curnow, G., 2002. Australian transport safety bureau heavy truck crash databases: what do the statistics tell us? [WWW Document]. Natl. Heavy Veh. Saf. Semin. http://www.ntc.gov.au/filemedia/Publications/WhatdoStatisticstellusGitaCurnow.pdf (accessed 2.09.13).

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.

Fatality Facts: Large Trucks, 2004. [WWW Document], 2004. URL http://images.businessweek.com/autos/pdfs/largetrucks.pdf (accessed 06.18.14).

Federal Highway Administration, 2013. Freight Facts and Figures 2011-Tables 2–1 and 2–1M. Weight of Shipments by Transportation Mode: 2007, 2010. and 2040 [WWW Document]. U.S. Dep. Transp. Fed. Highw. Adm. Off. Freight Manag. Oper. Freight Anal. Framew. URL http://www.ops.fhwa.dot.gov/freight/freight_analysis/nat_freight_stats/docs/11factsfigures/table2_1.htm.

Ghariani, N., 2001. Measures to Improve Truck Traffic Safety in Texas Roadways. ETD Collection for University of Texas, El Paso.

Gkritza, K., Mannering, F.L., 2008. Mixed logit analysis of safety-belt use in single- and multi-occupant vehicles. Accid. Anal. Prev. 40, 443–451.

Halton, J.H., 1960. On the efficiency of certain quasi-random sequences of points in evaluating multi-dimensional integrals. Numer. Math. 2, 84–90.

Islam, M.B., 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, 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. 67C, 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.
Kilcarr, S., 2014. Calculating the cost of crashes [WWW Document].
Knipling, R.R., Bocanegra, J., 2008. Comparison of Combination-Unit Truck and Single-Unit Truck Statistics from the LTCCS, FMCSA and Volpe Center Project report, No. DTRS57-04-D-30043.
Lemp, J.D., Kockelman, K.M., Unnikrishnan, A., 2011. Analysis of largetruck crash severity using heteroskedastic ordered probit models. Accid. Anal. Prev. 43, 370–380.
Martensen, H., Dupont, E., 2013. Comparing single vehicle and multivehicle fatal road crashes: a joint analysis of road conditions, time variables and driver characteristics. Accid. Anal. Prev. 60, 466–471.
McCartt, T.A., Rohrbaugh, J.W., Hammer, M.C., Fuller, S.Z., 2000. Factors associated with falling asleep at the wheel among long-distance truck drivers. Accid. Anal. Prev. 32, 493–504.
McFadden, D., 1981. Econometric models of probabilistic choice. In: Manski, C.F., McFadden, D. (Eds.), A Structural Analysis of Discrete Data with Econometric Applications. The MIT Press, Cambridge, MA, pp. 198–272.
Mcfadden, D., Train, K., 2000. Mixed MNL models for discrete response. J. Appl. Econom. 15, 447–470.
Morgan, A., Mannering, F.L., 2011. The effects of road-surface conditions age, and gender on driver-injury severities. Accid. Anal. Prev. 43, 1852–1863.
Offei, E., Young, R., 2014. Quantifying the Impact of Large Percent Trucks Proportion on Rural Freeways. North Dakota State University – Upper Greate Plains Transportation Institute, Mountains–Plains Consortium, Fargo.
Qin, X., Ivan, J.N., Ravishanker, N., Liu, J., Tepas, D., 2006. Bayesian estimation of hourly exposure functions by crash type and time of day. Accid. Anal. Prev. 38, 1071–1080.
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.
Shaheed, M.S., Gkritza, K., 2014. A latent class analysis of single-vehicle motocycle crash severity outcomes. Anal. Methods Accid. Res. 2, 30–38.
Train, K., 1999. Halton sequences for mixed logit. Berkley, CA, 1–18 (accessed from https://escholarship.org/uc/item/6zs694tp).
Train, K., 2003. Discrete Choice Methods with Simulation. Cambridge University Press, Cambridge, UK.
Washington, S., Karlaftis, M., Mannering, F., 2010. Statistical and Econometric Methods for Transportation Data analysis, second ed. Chapman and Hall/CRC, Boca Raton, FL.
Xiong, Y., Mannering, F.L., 2013. The heterogeneous effects of guardian supervision on adolescent driver-injury severities: a finite-mixture random-parameters approach. Transp. Res. Part B Method. 49, 39–54.
Zaloshnja, E., Miller, T.R., 2004. Costs of large truck-involved crashes in the United States. Accid. Anal. Prev. 36, 801–808.
Zaloshnja, E., Miller, T., 2006. Unit costs of medium and heavy truck crashes. Federal Motor Carrier Safety Administration, Washington, DC.
Zhu, X., Srinivasan, S., 2011. A comprehensive analysis of factors influencing the injury severity of large-truck crashes. Accid. Anal. Prev 43, 49–57.