Accident Analysis and Prevention 99 (2017) 249–261

Contents lists available at ScienceDirect

Accident Analysis and Prevention

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

Truck crash severity in New York city: An investigation of the spatial and the time of day effects

Zou Wei[a], Wang Xiaokun[b,*], Zhang Dapeng[a]

[a] Rensselaer Polytechnic Institute, 4027 JEC Building, 110 8th Street, Troy, NY 12180-3590, USA
[b] Department of Civil and Environmental Engineering, Rensselaer Polytechnic Institute, 4032 JEC Building, 110 8th Street, Troy, NY 12180-3590, USA

ARTICLE INFO

Article history:
Received 14 July 2016
Received in revised form
21 November 2016
Accepted 29 November 2016
Available online 13 December 2016

Keywords:
Truck crash severity
Vehicle weight
Spatiotemporal effect
Urban environment
Random parameter ordered probit model
Spatial generalized ordered probit model

ABSTRACT

This paper investigates the differences between single-vehicle and multi-vehicle truck crashes in New York City. The random parameter models take into account the time of day effect, the heterogeneous truck weight effect and other influencing factors such as crash characteristics, driver and vehicle characteristics, built environment factors and traffic volume attributes. Based on the results from the co-location quotient analysis, a spatial generalized ordered probit model is further developed to investigate the potential spatial dependency among single-vehicle truck crashes. The sample is drawn from the state maintained incident data, the publicly available Smart Location Data, and the BEST Practices Model (BPM) data from 2008 to 2012. The result shows that there exists a substantial difference between factors influencing single-vehicle and multi-vehicle truck crash severity. It also suggests that heterogeneity does exist in the truck weight, and it behaves differently in single-vehicle and multi-vehicle truck crashes. Furthermore, individual truck crashes are proved to be spatially dependent events for both single and multi-vehicle crashes. Last but not least, significant time of day effects were found for PM and night time slots, crashes that occurred in the afternoons and at nights were less severe in single-vehicle crashes, but more severe in multi-vehicle crashes.

1. Introduction

1.1. Background and motivation

Freight transportation is a critical part of the transportation system, especially for big cities. According to the 2012 Urban Mobility Report, the Greater New York Area ranked the first in total travel delay hours in 2011 (544 million hours), of which 20% were generated by trucks, resulting in a total congestion cost of $11.8 million (Schrank et al., 2012; Wang and Kockelman, 2005a,b). The New York Metropolitan Transportation Council (NYMTC) predicted that the volume of freight moving through the area is expected to increase 48% by 2040 (NYMTC, 2013). While the rapid growth in the freight industry stimulates the economic growth and provides more convenience to people's daily lives, it carries safety concerns. According to the Federal Motor Carrier Safety Administration (2013), 2757 people were killed and 88,000 people were injured in large truck crashes in 2012. The crashes bring both emotional burdens and economic losses to victims and the society, thus underscoring the importance of truck crash studies.

In response to the need to address transportation safety concerns, a wide variety of modeling techniques have been developed in the literature to study the contributing factors of the road crashes. For comprehensive reviews, see Lord and Mannering (2010) for crash frequency analysis, Savolainen et al. (2011) for injury severity analysis, and Mannering and Bhat (2014) for more advanced models. Among all kinds of road crashes, a crash involving trucks is a complex event with unique characteristics, and should be studied separately. Similarly, single-vehicle and multi-vehicle crashes usually have different causes; therefore, using different models can help to identify the confounding factors more easily.

Besides the differentiation in vehicle types, spatial and temporal dependence among neighboring segments are also ignored in many crash modeling studies. In late 2009, the New York City Department of Transportation has implemented an "Off-Hour Truck Delivery Program" to encourage truckers to shift their deliveries to off-hours (between 7:00 p.m. and 6:00 a.m.) (Holguín-Veras et al., 2010a,b). The success of the Off-Hour Delivery program is widely recognized, and the United States' Federal Highway Administration has decided to apply the concept nationwide (Federal Highway Administration, 2012). The safety concerns of the program, for example, the light-

* Corresponding author.
E-mail addresses: zouw2@rpi.edu (W. Zou), wangx18@rpi.edu, wangxiaokun@gmail.com (X. Wang), zhangd9@rpi.edu (D. Zhang).

http://dx.doi.org/10.1016/j.aap.2016.11.024
0001-4575/© 2016 Published by Elsevier Ltd.

ing condition and drivers' fatigue at night, are raised at the same time. Although the original study the off-hour delivery program has shown that truckers were less stressful when driving at night (Holguín-Veras et al., 2010a,b), the general application validity is limited by the small study size. To fully understand the safety issues, the temporal effect should be considered in the truck crash analysis. Similarly, the spatial dependency effect should also be considered in truck crash analysis in big cities, where the uncontrolled factors, such as the pedestrian volume, tend to be similar between closer crash sites due to the compact urban form. In terms of modeling accuracy, ignoring the spatial and temporal dependence violates the sample independence assumption, and could result in inconsistent and confounding estimates, and efficiency losses (Aguero-Valverde and Jovanis, 2008, 2010; Chiou and Fu, 2015; Cressie, 2015; Dubin, 1988; LeSage, 2009; LeSage and Pace, 2009; Lord and Mannering, 2010; Lord and Persaud, 2000; Ni et al., 2016; Zhang and Wang, 2014; Zhang and Wang, 2015a,b; Zhang and Wang, 2016).

Last decade's "SUV boom" has initiated the discussion on the effect of vehicle weight on crash severity (Hakim, 2004). Literature has shown that heavier/larger vehicles tended to provide more protection to their drivers against fatalities (Bedard et al., 2002; Kahane, 2003), however, the opposite might be true for occupants in their collision partners (White, 2003). When it comes to freight transportation, the vehicle weight requires even more attention when safety concerns are addressed. There are multiples cities in the US and Europe that have banned/restricted the entrance of heavy trucks in the city to avoid potential crashes. Therefore, it is important to explore whether vehicle weight has a heterogeneity effect on the injury severity of truck crashes, so that in the future, it may help planners to propose more relevant road safety policies for freight deliveries, and help engineers to develop better vehicle designs for trucks.

The objective of this paper is to analyze the influencing factors of crash severity for both single and multiple vehicle truck crashes that occurred in New York City from 2008 to 2012, recognizing three major important issues understudied in traditional crash severity analysis: 1) the time of day effect 2) the spatial dependency effect, and 3) the heterogeneous effect of truck weight. To do so, two random parameter ordered probit models (Zhang et al., 2014) are applied to analyze the connections among the New York City's truck crashes, and the potential contributing factors of these crashes. A spatial generalized ordered probit model is used to further investigate the spatial dependency effect among single and multi-vehicle truck crashes. An integration of the state-maintained incident data, the publicly available smart location data, and the BEST Practices Model (BPM) data allows the examination of a wide range of factors including the crash, driver and vehicle characteristics, the traffic volume on roads, and the built environment attributes.

1.2. Literature review

1.2.1. Sub-group crash modeling

In recent years, researchers have proposed the use of sub-group crash models to distinguish the different characteristics associated with crashes from different categories. Two most popular classifications are related to the number of vehicles involved in an crash and the vehicle type. Numerous studies have shown that single-vehicle and multi-vehicle crashes have vastly different exposure and geometric design feature attributes, and distinct models should be developed to account for such differences (Chen and Chen, 2011; Geedipally and Lord, 2010; Griffith, 1999; Ivan, 2004; Ivan et al., 1999, 2000; Kockelman and Kweon, 2002; Lord et al., 2005; Mensah and Hauer, 1998; Öström and Eriksson, 1993; Shankar et al., 1995). For more recent work, Wu et al. (2014) developed mixed logit models to analyze driver injury severities in single-vehicle and multi-vehicle crashes on rural two-lane highways. The results indicated that drivers had more severe injuries in multi-vehicle crashes when motorcycles or trucks were involved, and when there were dark lighting conditions or dusty weather conditions; drivers had higher probability of having severe injuries in single-vehicle crashes when vans were involved and drivers' overtaking actions were identified in the crash. Yu and Abdel-Aty (2013) used Bayesian models to identify two different sets of significant explanatory and exposure variables for single- vehicle and multi-vehicle crashes. The authors found that although both multi-vehicle and single-vehicle crash occurrences were associated with road design features such as the number of lanes, degrees of curvatures and median widths, multi-vehicle crashes were also related to curve length ratios and segment length and single-vehicle crashes were more relevant to speed limits and longitudinal grades. Martensen and Dupont (2013) compared single-vehicle and multi-vehicle fatal crashes in six European countries and found that the traffic, the presence of a juction/physical division between carriageways were the most important variables to distinguish these two classes of crashes.

Similary, distinct crash prediction models have been developed for different types of vehicles in the literature, especially for passenger vehicles and trucks (Jovanis and Chang, 1986; Lee and Abdel-Aty, 2005; Miller et al., 1998). For truck crash analysis, existing literature focuses heavily on injury severity analysis for individual large trucks on highways. Khattak et al., (2003) used ordered probit models to analyze the truck driver's injury severity in large truck crashes in North Carolina, from 1996 to 1998, using the Highway Safety Information System data. The authors found that roll-over crashes tend to generate more severe injuries in single truck crashes, and driver's behavior such as drug use and speeding also increased the crash severity. Golob et al. (1987) used the Traffic Crash Surveillance and Analysis data to assess the truck-involved freeway crashes severity in Los Angeles area from 1983 to 1984. The result showed that "hit-object" and "rear end" were the most dangerous types of incidents. Zhu and Srinivasan (2011) used the Large Truck Crash Causation Study (LTCCS) data to study the influencing factors of large truck crash severity in 17 states in the US, from April 2001 to December 2003. The estimates from ordered logit models captured the negative impacts of driver behavior, such as truck driver distraction, alcohol use and emotional factors, on crash severity. Lemp et al. (2011) found that crash severity increased with the number of trailers, but fell with the truck length and gross vehicle weight rating. Islam and Hernandez (2013a,b,c) used both a random parameter ordered probit model and a mixed logit model to analyze the injury severities of multi-vehicle collisions involving large trucks, using a fused national crash dataset. The result shows that the level of injury severity is a result of complex interaction of human factors, such as distracted/sleepy driving, female occupants, and seat-belt usage; road and environmental facts such as light conditions; road geometries such as curved segments and wet surface; vehicle characteristics such as vehicle and traffic conditions. The authors also used a mixed logit model to study the large truck crash injury outcomes in Texas, revealing that the complex interactions between factors including driver demographics, traffic flow, roadway geometric features, land use, time characteristics, weather and light conditions, contributed to the different level of injury outcomes (Islam and Hernandez, 2013b).

1.2.2. Spatial dependency, temporal dependency and heterogeneity analysis

The subjective selection of sample data (temporal and spatial segmentation) for crash studies may result in potential spatial and temporal dependence among observations. Neighboring sites typically have similar environmental and geographical characteristics and may share unobserved effects. For example, road

Table 1
The Summary Statistics for Single-vehicle and Multi-vehicle Crashes (Continuous Variables).

Single-vehicle Crashes Attributes	Descriptive statistics		Minimum	Maximum
	Mean	Standard Deviation		
Mean truck weight	34,168	25,046	3380	160,000
Mean driver's age	42	11	17	83
Total hourly traffic volume	934	1142	0	11,165
Single occupancy vehicle hourly traffic volume	511	678	0	6798
High occupancy vehicle hourly traffic volume	184	237	0	1572
Taxi hourly traffic volume	41	73	0	691
Truck hourly traffic volume	106	181	0	1306
Commercial van hourly traffic volume	76.8	124	0	1541
Bus hourly traffic volume	15.1	27.7	0	327
Retail business establishment density (jobs/acre)	0.566	2.502	0	40.4
Office establishment density (jobs/acre)	1.248	10.2	0	123
Industrial business establishment density (jobs/acre)	1.02	4.25	0	41.1
Service business establishment density (jobs/acre)	0.571	2.05	0	32.8
Entertainment business establishment density (jobs/acre)	0.406	1.38	0	15.6
Multi-vehicle Crashes Attributes	**Descriptive statistics**		**Minimum**	**Maximum**
	Mean	Standard Deviation		
Total hourly traffic volume	1387	1518	0	11,165
Single occupancy vehicle hourly traffic volume	823	966	0	6798
High occupancy vehicle hourly traffic volume	293	334	0	1599
Taxi hourly traffic volume	32.5	50.3	0	486
Truck hourly traffic volume	151	222	0	1335
Commercial van hourly traffic volume	73.3	98.4	0	1541
Bus hourly traffic volume	13.2	28.9	0	289
Retail business establishment density (jobs/acre)	0.572	3.15	0	58.3
Office establishment density(jobs/acre)	1.063	9.84	0	162.8
Industrial business establishment density(jobs/acre)	1.084	6.74	0	128.4
Entertainment business establishment density(jobs/acre)	2.11	32.8	0	714.1
Mean truck weight	22,305	13,871	2193	99,500
Maximum truck weight	41,315	27,294	2478	160,000
Minimum truck weight	4125	3920	531	92,000
Mean driver's age	42	9	19	82

design features influencing crash occurrences at one site may have unobserved "spatial spillover" effects to other sites located proximally. At the same time, the variates affecting crash occurrences may change (e.g., the lighting condition) or stay consistent (e.g., road conditions) depending on the time of day. Recently, numerous efforts have been spared to include spatial correlation in the development of collision models (see Wang et al. (2012) for a comprehensive review), for both area-wide (Aguero-Valverde, 2013; Aguero-Valverde and Jovanis, 2008, 2006; Amoros et al. 2003; Flask and Schneider, 2013; Guo et al., 2010; MacNab, 2004; Noland and Quddus, 2004; Quddus, 2008) and road segments/intersections (Abdel-Aty and Wang, 2006; Aguero-Valverde and Jovanis, 2008, 2010; El-Basyouny and Sayed, 2009; Guo et al., 2010; Mitra, 2009).

As to studies accommodating time of day effect/temporal dependence, indicator variables are usually used and employed in the models (Chang and Chen, 2005; Dupont et al., 2013; Islam and Hernandez, 2013a,b; Qin et al., 2006). However, such modeling approach does not account for the heterogeneity in variable effects during different time periods (Pahukula et al., 2015), and a possible solution is to estimate separate univariate count models for each time period (Mannering and Bhat, 2014). Recently Pahukula et al. (2015) studied the time of days effect for crashes involving large trucks occurring on urban freeways in Texas, from year 2006 to 2010. The results of five several random parameter models of different time of the day have demonstrated the traffic flow, light conditions, surface conditions, time of year and percentage of trucks on the road influence the injury severity differently during different times of the day. However, since the unobserved factors across different time periods are likely to affect multiple dependent variables (i.e., crash counts at different times of day) simultaneously (Mannering and Bhat, 2014), the separated univariate count models are not the optimum approach. A number of studies also began to explore the spatial and temporal correlations simultaneously. Castro et al. (2012) proposed a latent variable-based generalized ordered response framework for count data models, in which spatial dependence is introduced by a spatial structure on the latent continuous variables, and temporal dependence is captured in the error term of the latent variable. Chiou and Fu (2015) developed a multinomial generalized Poisson model with error components and spatiotemporal dependence (ST-EMGP) that not only models crash count and injury severity simultaneously, but also accommodates spatial and temporal dependence. The above studies focus on a longer temporal segment (in years); however, the time of day effect is not studied. Zou et al. (2015) studied the time of day effect of truck crashes in New York City, however, the results are limited by the small sample size and modeling techniques, which leads to the development of this paper, with more complete data and more comprehensive modeling designs.

Another fundamental issue that requires more attention in the crash modeling literature is the presence of the unobserved heterogeneity, which may otherwise seriously bias the model estimates and provide invalid statistical tests (Karlaftis and Tarko, 1998). Relatively recent research has demonstrated the importance of considering the heterogeneity across roadway and traffic conditions, vehicle characteristics, drivers' characteristics, collision types and crash-specific details using various modeling approaches (Barua et al., 2016); Milton et al. (2008), Eluru and Bhat (2007), Eluru et al. (2008), Eluru et al. (2010), Kim et al. (2013), Paleti et al. (2010), Venkataraman et al. (2011), Morgan and Mannering (2011), Aziz et al., (2013), Zou et al. (2013), Yasmin and Eluru (2013), Bhat et al. (2014), Yasmin et al. (2014). While the heterogeneity effect of vehicle weight on road crashes has been studied in a number of papers (Bedard et al., 2002; Lui et al., 1989; Miller et al., 1999; Wang and

Kockelman, 2005a,b; Zou et al., 2015), it is a topic that has not been explored fully in the truck crash analysis.

This paper aims to contribute to the literature by conducting a comprehensive analysis on both single-vehicle and multi-vehicle truck crashes in New York City, accounting for the time of day effect, the unobserved heterogeneity weight effect and the spatial dependency effect. In doing so, two distinct random parameter ordered probit models are developed to analyze the influencing factors of single-vehicle and multi-vehicle truck crashes separately. Based on the spatial auto-correlation analysis, two spatial generalized ordered probit models are used to analyze the spatial dependency effects in single-vehicle and multi-vehicle truck crashes. To the authors' best knowledge, this is the first study using econometric models that account for temporal effects, spatial effects and unobserved heterogeneity to analyze the truck crashes in a big city.

The rest of the paper is structured as follows. The next section presents the data sources and data processing procedures. Then the modeling framework is described, and the empirical estimation results are presented. The final section summarizes the findings and concludes the study.

2. Data description

2.1. Data collection and processing

New York City is chosen as the study area due to its diverse demographics, land use, drivers' behavior and the high demand of freight deliveries. As the center for business and tourism, the city has the highest population density in the US, which generates huge demand of resources such as food, packages and household items. As a result, thousands of trucks are traveling in the city to make deliveries every day, which causes transportation conflicts such as congestion and crashes.

The three primary data sources used in this study are: the state maintained incident data, the publicly available smart location data, and the BEST Practices Model (BPM) data. The integration of these three datasets allows the examination of a wide range of factors comprehensively, including crash characteristic, vehicle and driver characteristics, environmental factors, built environment factors, and traffic information.

The original incident data is provided by the New York State Department of Transportation (from year 2008–2012) with three separate files: the event information, the vehicles involved, and the contributing factors (NYSDOT, 2014).

The smart location data is developed and maintained by the Environmental Protection Agency (EPA) in 2013. It includes various built environment information such as land use, urban design, demographics and transit services, generated from data sources including Census of Population and Economic Census, and American Community Survey. The majority of the information was collected in 2010 (US EPA, 2014).

The BPM data is developed by the New York Metropolitan Transportation Council (NYMTC) with road traffic volume during four time blocks: (1) A.M., from 6:00 to 10:00 A.M., (2) Midday (MD), from 10:00 A.M. to 3:00 P.M., (3) P.M., from 3:00 to 7:00 P.M., and (4) Night (NT), from 7:00 PM to 6:00 A.M. To ensure the consistency in modeling estimation, hourly traffic volume is derived for each time block. Since the BPM data does not include traffic information on smaller links (Brinckerhoff, 2009), the average traffic flow on the two nearest links of an crash was used for the modeling.

For the current study, truck crashes occurring in New York City from 2008 to 2012 were extracted from the state maintained incident data with a focus on truck crashes. A truck crash is defined as an incident involving at least one truck. The types of trucks analyzed in this study include: small wheel truck, delivery truck, dump truck, flatbed truck, stake truck, tank truck, refrigerated truck, tow truck, van truck, utility truck, and other unknown trucks, based on the classification of vehicle body type code provided in the NYS Incident Vehicle Data File (NYSDOT, 2014). The crash record includes single-vehicle crashes and multi-vehicle crashes. As New York City has a very high population density, many truck crashes involved pedestrians and bikers. These crashes are also included in this study, and they are flagged separately in the model estimates (i.e., by using indicator variables) to recognize the unique characteristics of pedestrian/biker-involved crashes.

The information was then compiled with the smart location data and the BPM data by using the "joint by location" function in QGIS (Gandhi, 2014). After filtering out attributes that are not relevant to this study, such as number of zero-car households per census block group, and attributes with incomplete information, a crash-based severity analysis dataset with 735 single-vehicle truck crashes and 3769 multi-vehicle truck crashes was used for model estimation. The original crash severity refers to the most severe injury to any person (including drivers, occupants) involved in a road vehicle crash (in or outside the vehicle) and is recorded in four ordinal categories: (1) property damage only (PDO), (2) non-incapacitating injury, (3) incapacitating injury, and (4) fatal. The next section discusses additional sample details on selected independent variables in this analysis.

2.2. Sample characteristics

Figs. 1 and 2 present the geographical distribution of single-vehicle and multi-vehicle truck crashes that occurred in five districts of New York City (Staten Island, Manhattan, Bronx, Queens, and Brooklyn) from 2008 to 2012, with their corresponding crash severity levels.

Fig. 1 shows that fewer single-vehicle crashes have occurred during the five years compared to multi-vehicle crashes, and they occurred more frequently in the Manhattan district. Out of all the single-vehicle crashes, the crash severity level was dominated by incapacitating injury crashes (66%), and the remaining ones were property non-incapacitating injury crashes (21%), property damage only crashes (6%), and fatal crashes (6%). The data show that single-vehicle crashes occurred more closely in groups in city centers with dense population, complex traffic condition, and urban forms.

Fig. 2 shows that multi-vehicle truck crashes occurred more frequently in Manhattan, Bronx, Brooklyn and upper Queens. Out of all the multi-vehicle crashes, around 37% of them had property damage only, 22% had non-incapacitating injuries, 41% had incapacitating injuries and only 0.4% of the crashes were fatal crashes. Compared to single-vehicle crashes, multi-vehicle crashes were distributed more evenly in all major areas of New York City, and also had a high percentage of incapacitating injury outcomes.

The combined dataset contains very detailed information related to truck crashes, including vehicle characteristics, drivers' characteristics, roadway geometry condition, weather condition, light condition, traffic volume, built environment factors, and time of day indicator variables. Tables 1 and 2 provide the summary statistics for both single-vehicle and multi-vehicle truck crashes.

For single-vehicle crashes, the average age of truck drivers was 42 years old with standard deviation of 11 years and the youngest was only 17 years old. The traffic volume data show that single occupancy vehicles accounted for the largest percent traffic volume in all the time blocks. For the business establishment densities, office has the highest density, followed by industrial business, and the retail, service and entertainment business densities were similar to each other. The average truck weight was ranging from 3380 to 160,000 pounds, with an average of 34,168 pounds.

As for multi-vehicle crashes, they had higher average single and high occupancy vehicle traffic volumes than single-vehicle crashes.

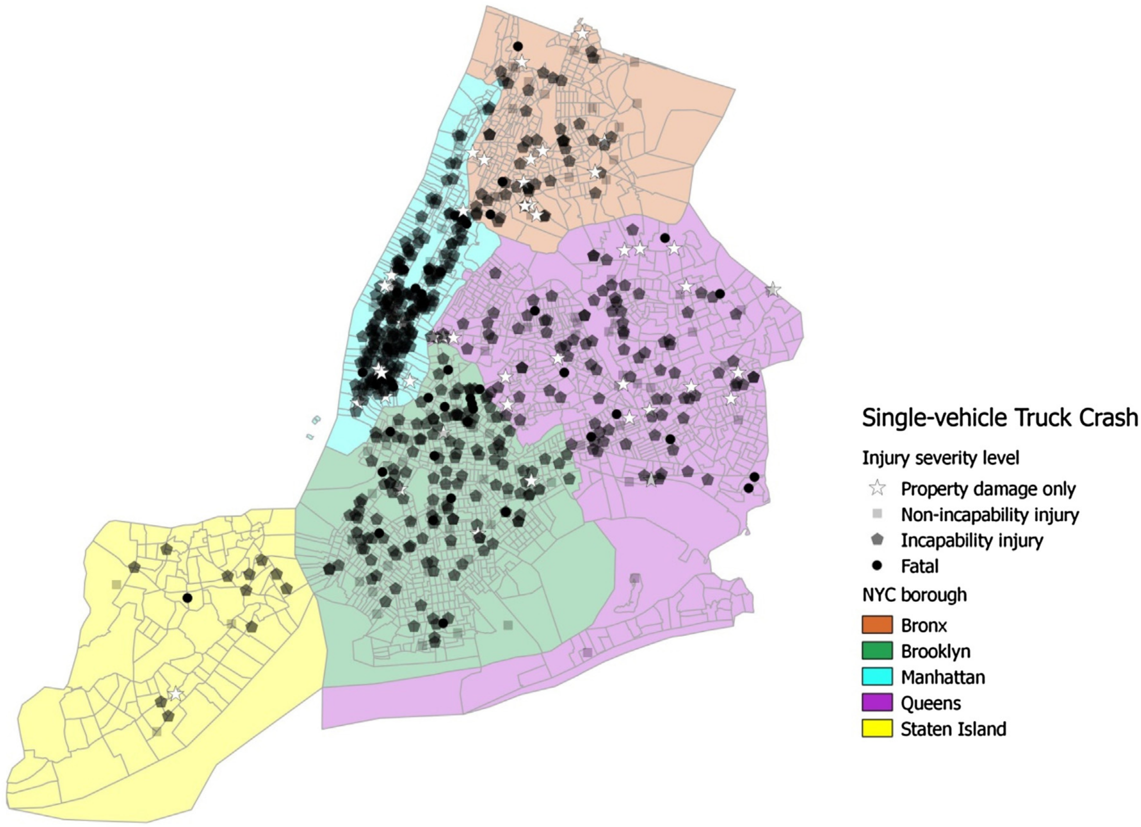

Fig. 1. Crash Severity Distribution of Single-vehicle Crashes.

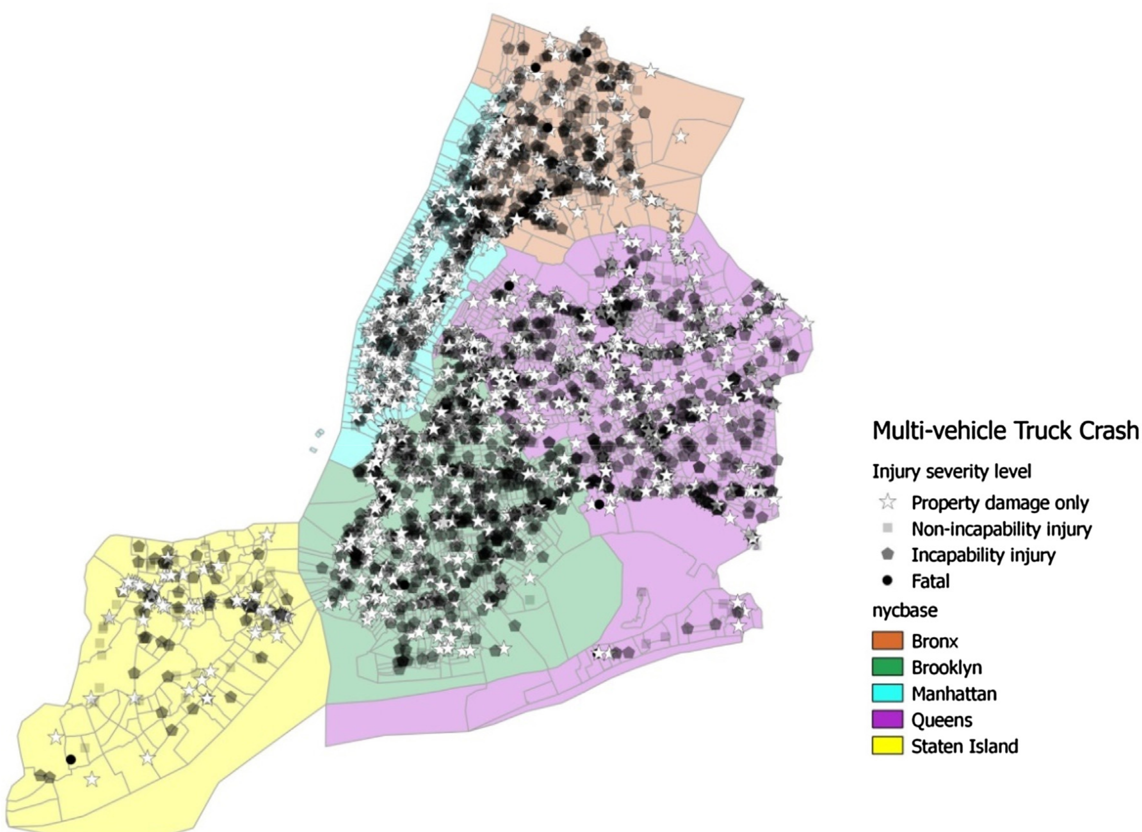

Fig. 2. Truck Crash Severity Distribution of Multi-vehicle Crashes.

Table 2
The Summary Statistics for Single-vehicle and Multi-vehicle Crashes (Categorical Variables).

Single-vehicle crashes	
Variable	Share
Injury severity level	
Incapacitating injury	66%
Non-incapacitating injury	22%
Property damage only	6%
Fatal	6%
Crash types	
Pedestrian crashes	60%
Biker crashes	21%
Other	19%
Traffic conditions	
Traffic signal	56%
None	38%
Stop sign	2%
Other	2%
Light conditions	
Daylight	74%
Dark road	22%
Dawn	2%
Dusk	2%
Weather	
Clear	72%
Cloudy	14%
Rain	12%
Snow	1.90%
Other	0.10%
Road geometries	
Straight and level	89%
Straight and grade	5%
Straight at hill crest	1%
Curve and level	2%
Curve and grade	2.50%
Curve and hill crest	0.50%
Road surface conditions	
Dry	80%
Wet	17%
Snow/Ice/Slush	2%
Other	1%
Time of day	
AM	22%
Mid-day	38%
PM	18%
Night	22%

Multi-vehicle crashes	
Variable	Share
Injury severity level	
Non-incapacitating injury	41%
Property damage only	37.60%
Incapacitating injury	22%
Fatal	0.40%
Crash types	
Pedestrian crashes	0.10%
Biker crashes	0.05%
Other	99.85%
Traffic conditions	
Traffic signal	45%
None	44%
Stop sign	6%
Other	3%
Yield sign	0.40%
Highway construction	1.20%
Maintenance area	0.40%
Light conditions	
Daylight	81%
Dark road	15%

Table 2 (*Continued*)

Multi-vehicle crashes	
Variable	Share
Dawn	2%
Dusk	2%
Weather	
Clear	72%
Cloudy	15%
Rain	11%
Snow	1%
Sleet	0.40%
Other	0.60%
Road geometries	
Straight and level	86%
Straight and grade	7%
Straight at hill crest	2%
Curve and level	2%
Curve and grade	2%
Curve and hill crest	1%
Road surface conditions	
Dry	81%
Wet	16%
Snow/Ice/Slush	2%
Other	1%
Time of day	
Mid-day	39%
AM	23%
PM	23%
Night	15%

For business establishment densities, entertainment business was the highest, followed by industrial and office establishments. As there might be multiple trucks involved in a multi-vehicle crash, the maximum, minimum and average truck weights are explored in the study. The average age for truck drivers was similar to the one for single-vehicle crashes.

For the categorical variables, the data shows that most single-vehicle crashes involved pedestrians/bikers. For traffic condition attributes, more than 50% of the crashes (both single and multi-vehicle) occurred when traffic signals were on. The light condition information shows that most crashes were occurred during daytime (70%–80%), and on dry surface roads (80%). For time of day attributes, most crashes occurred during the Mid-day period (for both single and multi-vehicle crashes), and there were fewer crashes occurred during the pm and night periods.

Figs. 3 and 4 present the distribution of truck registered-weight in single-vehicle crashes and multi-vehicle crashes respectively. The trucks' registered-weights are divided into 10 groups. Fig. 3 shows that the distribution of truck registered weights in all crash severity levels varies widely. In multi-vehicle crashes, Fig. 4 shows that lighter trucks have a higher rate of more severe crash injury levels.

2.3. Spatial auto-correlation analysis

The colocation quotient (CLQ) is a metric used to measure the spatial association between pairs of crashes of different injury severity levels, given the clustering geometric pattern of crash occurrences (Leslie and Kronenfeld, 2011). The CLQ of injury level i with respect to injury level $\left(\mathrm{CLQ}_{i\to j}\right)$. is formally defined as:

$$CLQ_{i\to j} = \frac{C_{i\to j}/N_i}{N_j^{'}/(N-1)}$$

where N = the total number of crashes of all injury levels

N_i = the number crashes with injury severity level i (i.e. type i crash)

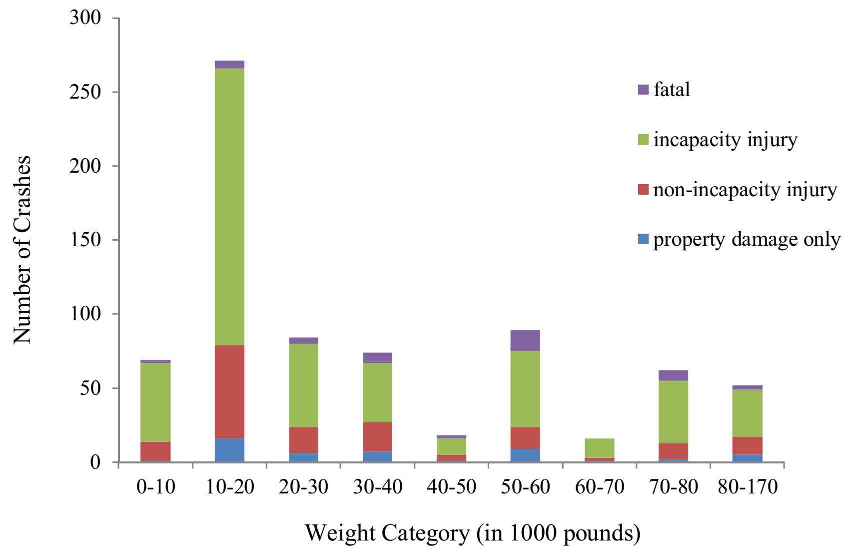

Fig. 3. Average Truck Weight and Crash Severity with Histograms (Single-vehicle).

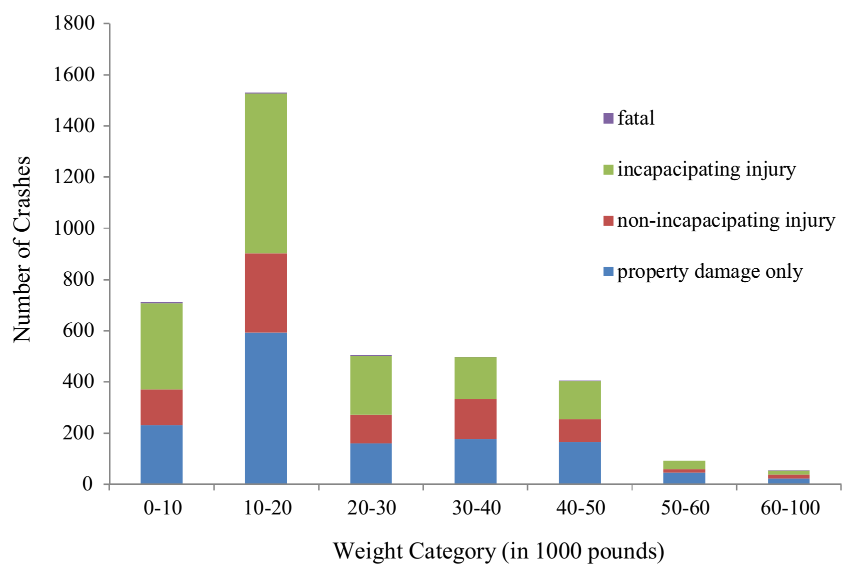

Fig. 4. Average Truck Weight and Crash Severity with Histograms (Multi-vehicle).

$N_j^{'}$ = the number of crashes with injury severity level j (minus 1 when $i=j$)

$C_{i\rightarrow j}$ = the number of crashes whose injury severity level is i and the nearest neighbor is an crash with injury severity level j.

The numerator of the formula is the proportion of type j crashes among type i crashes' nearest neighbors (i.e., the observed proportion), and the denominator is the proportion of type j crashes that could be a nearest neighbor to each type i crash (i.e., the expected proportion). As a result, the colocation quotient between i and j measures the degree to which type j crash attracts type i crashes, comparing to the random chances. For example, $CLQ_{i\rightarrow j} = 1.5$ indicates that a type i crash is 1.5 as likely to have a type j crash as its nearest neighbor, comparing to by chance. The definition of the colocation quotient is naturally unidirectional (i.e., $CLQ_{i\rightarrow j}$ not necessarily equal to $CLQ_{j\rightarrow i}$). When the nearest neighbor relationship is asymmetric, for example, when i's nearest neighbor is j, but j's nearest neighbor is not i, then $CLQ_{i\rightarrow j}$ will be larger than $CLQ_{j\rightarrow i}$.

To explore the spatial autocorrelations in both single-vehicle and multi-vehicle crashes, the local pair-wise colocation quotients are estimated and summarized in Table 3. For single-vehicle crashes, the same-category CLQs tend to be higher for property damage only and fatal crashes. For example, the crashes involving PROPERTY DAMAGE ONLY were 1.79 times more likely to have neighbor crashes that were PROPERTY DAMAGE ONLY as well, and the fatal crashes were 1.55 more likely to have fatal crashes as their nearest neighbors. For the category to CLQs, the PROPERTY DAMAGE ONLY crashes were more likely to have non-incapacitating crashes nearby (CLQ = 1.45), and the non-incapacitating crashes tend to have more incapacitating crashes as their neighbors (CLQ = 1.08). The other two types of crashes (Incapacitating and fatal crashes) show similar trends. For multi-vehicle crashes, the same-category colocation quotient indicates that crashes involving PROPERTY DAMAGE ONLY were 1.14 times more likely to have neighbor crashes that were PROPERTY DAMAGE ONLY as well. Similarly, crashes involving non-incapacitating injury were 1.03 times more likely to have neighbor crashes that had an injury as well. The Fatal-Fatal colocation quotient is zero in multi-vehicle crashes, indicating that it is very unlikely that a fatal crash would occur "next to" another fatal crash. This phenomenon may be due to the fact that there are only a few cases of fatal crashes in the multi-vehicle

Table 3
The Colocation Quotients of Single-vehicle and Multi-vehicle Crashes.

	Property Damage Only	Non-incapacitating Injury	Incapacitating Injury	Fatal
Single-vehicle (N = 735)				
Property Damage Only	1.789	1.452	0.836	0.348
Non-Incapacitating Injury	0.823	0.902	1.078	0.686
Incapacitating Injury	0.977	0.833	1.049	1.083
Fatal	1.043	0.528	1.101	1.552
Multi-vehicle (N = 3769)				
Property Damage Only	1.135	0.951	0.909	0.870
Non- Incapacitating Injury	0.948	1.031	1.030	0.902
Incapacitating Injury	0.898	1.033	1.074	0.928
Fatal	1.0	0.894	1.066	0

crash dataset, which makes it difficult to identify the neighboring effect.

Based on the above results, single-vehicle crashes tend to have strong spatial autocorrelations as their CLQs deviate substantially from one. In contrast, the CLQs for multi-vehicle crashes imply that the spatial autocorrelations among multi-vehicle crashes seems to be less significant. Spatial generalized ordered probit models are estimated in later sections to further explore the spatial dependency effects amongst crashes.

3. Methodology

The corresponding model specification starts from a standard ordered probability model, which captures the influence of independent factors on the ordinal ranking outcomes. The ordinal feature is derived from a latent variable, which is typically specified as a linear function for each truck crash, such that:

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

where X is a set of independent variables (factors influencing crash severity in this study), $\boldsymbol{\beta}$ is the corresponding estimable parameters, and ε is the set of random disturbance terms representing the unobserved effects and errors.

The actually observed responses (i.e. the injury severity level), y is defined as follows:

$y = 0$ if $-\infty \le y^* \le \alpha_1$ (Property damage only)

$y = 1$ if $\alpha_1 < y^* \le \alpha_2 \alpha_1 < y^* \le \alpha_2$ (Non-incapacitating Injury)

$y = 2$ if $\alpha_2 < y^* \le \alpha_3$ (Incapacitating Injury)

$y = 3$ if $y^* > \alpha_3$ (Fatal)where α set of cut-points that define the integer ordinal outcomes. When disturbance term ε is assumed to be normally distributed, the model becomes an ordered probit model.

A random parameter ordered probit model allows regression parameters to vary across crash locations according to a specified distribution. Let q indicates the index of each truck crash, the latent variable can be specified as:

$$y_q^* = \boldsymbol{X}_q \boldsymbol{\beta}_q + \varepsilon \tag{2}$$

where the random parameter $\boldsymbol{\beta_q}$ represents the unique characteristics of individual q. The random parameter ordered probit model is estimated using the simulated maximum likelihood, which is embedded in the RChoice package developed by Sarrias (2014).

A generalized ordered probit model adds flexibility to a normal ordered probit model by allowing the thresholds (αs) to vary across individuals (Eluru et al., 2008):

$$\alpha_{q,k} = \alpha_{q,k-1} + \exp\left(\tau_k + \boldsymbol{\gamma}_k \boldsymbol{z}_{qk}\right) \tag{3}$$

where τ_k is a scalar, $\boldsymbol{\gamma}_k$ is a vector of coefficients associated with injury severity level k. For identification reasons, $\boldsymbol{\gamma}_1$ is normalized to zero for $\alpha_{q,1}$ (i.e., $\alpha_{q,1} = \exp(\tau_1)$).

To take into account the spatial dependency among the truck crashes, a spatial lag generalized ordered probit model (Castro et al., 2013a,b) is used. A predefined spatial proximity matrix W, which indicates the relative distances between observations is introduced in the formulation of the latent injury severity variable:

$$y_q{}^* = \rho \sum_{q'=1}^{Q} w_{qq'} y_{q'}{}^* + \boldsymbol{X}_q \boldsymbol{\beta}_q + \varepsilon_q \tag{4}$$

Where ρ is the spatial autocorrelation coefficient ranging from 0 to 1, $\boldsymbol{X}_q$ is a $(1 \times K)$ vector of variables that corresponds to observation q, $\boldsymbol{\beta}_q$ is a $(K \times 1)$. vector of coefficients. The row-standardized spatial proximity matrix W could be defined by several different approaches. Here, the distance between crash locations was used as a measure of spatial proximity, which was estimated from the UTM coordinate informion (in meters) provided by the state maintained crash incident dataset. A numr of functional forms were considered for the weight matrix, including different orders of inverse distances. The inverse of the cubed distance specification provided the best fit and was used for the final model specification. Different distance bands were also explored for the pairs of observations (1, 5, 10 miles), and 10 miles was selected to ensure that all crashes have at least three neighbors (i.e., crashes). The model goodness of fit also proves to be the best with the 10-mile band.

In order to deal with the computational burden caused by estimating the q-dimensional integral in the likelihood function, the pairwise composite marginal likelihood (CML) method is used to estimate the parameters (Castro et al., 2013a,b).

4. Result and discussion

4.1. Variable specification

The sample data were first separated into two datasets: a data set for single-vehicle crashes, with 735 observations at 689 locations, and a dataset for multi-vehicle crashes, with 3769 observations at 3085 locations. For both datasets, the methodological framework is formed by first identifying variables that can be expected to be explanatory when predicting crash severity and variables that are likely to capture any remaining unobserved effects. Initially, more than 40 independent variables were selected based on previous research and intuitive considerations. To obtain the final variable specification, extensive specification testing and correlation checking were carried out, and statistically insignificant variables were dropped step by step. Some categorical exogenous variables were re-grouped to ensure a sufficient number of observations within each level and levels showing similar effects were combined into one level. For random parameter analysis, all statistically significant explanatory variables were selected as random parameters at first, with various distribution forms (uniform, normal, triangular and lognormal). In the end, the random coefficient of “Truck Registered Weight” (normally distributed) was added to the final models based on statistical significance and the purpose of this study. The final specification and estimation results

Table 4
Random Parameter Ordered Probit Model Estimation Results (Single-vehicle Crashes).

Types	Variables	Coef.	s.d	t-stat	p-value	Marginal Effects			
						Property damage only	Non-incapacitating Injury	Incapacitating Injury	Fatal
Vehicle Characteristics	Truck registered weight_Mean	0.009	0.002	5.15	<0.001	−5e-04	−0.002	0.002	5e-04
	Truck registered weight_Variance	0.01	0.001	8.71	<0.001				
Crash Characteristics	Pedestrian crash	2.02	0.137	14.81	<0.001	−0.421	0.509	0.127	−0.215
	Biker crash	1.81	0.160	11.31	<0.001	−0.052	−0.312	0.007	0.296
Road Geometries (base: Straight and level)	Curve and grade	0.854	0.319	2.68	0.007	−0.217	−0.168	0.079	0.109
Traffic Volume	High occupancy vehicles	0.001	0.0002	2.58	0.009	−3e-05	−1e-04	1e-04	3e-05
Number of observations: 735	Threshold 1: 1.32 Threshold 2: 3.94 Log-likelihood: −597.18 AIC: 1210.4								

Table 5
Random Parameter Ordered Probit Model Estimation Results (Multi-vehicle Crashes).

Types	Variables	Coef.	s.d	t-stat	p-value	Marginal Effects			
						Property damage only	Non-incapacitating Injury	Incapacitating Injury	Fatal
Vehicle characteristics	Truck registered weight difference_Mean	0.002	6e-04	2.88	0.005	−6e-04	−0.13e-05	6e-04	1.6e-05
	Truck registered weight difference_Variance	2e-04	4e-04	0.58	0.562				
Road geometries (base: Straight and level)	Curve and grade	−0.283	0.144	−1.96	0.050	0.111	−0.008	−0.102	−0.002
	Curve and hill crest	1.12	0.434	2.57	0.010	−0.310	−0.097	0.357	0.051
Traffic volume	High occupancy vehicles	2e-04	0.6e-04	3.27	0.001	0.16e-06	0.75e-04	0.2e-05	−7.7e-05
	Taxi	−0.001	4e-04	−3.26	0.001	5e-04	0.11e-05	−5e-04	−1.3e-05
Business Establishment Density	Office	−0.013	0.004	−2.99	0.003	0.005	0.11e-04	−0.005	−1.3e-05
	Industrial	0.022	0.008	2.80	0.005	−0.009	−0.18e-04	0.008	2.2e-04
	Entertainment	0.071	0.188	3.77	<0.001	−0.027	−0.58e-04	0.027	0.001
Number of observations: 3769	Threshold1: 0.541 Threshold 2: 2.994 Log-likelihood: −4084.7 AIC: 8203.6								

Table 6
Spatial Generalized Ordered Probit Model Estimation Results (Single-vehicle Crashes).

Variables	Latent injury risk propensity		Threshold between property damage only and non-incapacitating injury		Threshold between non-incapacitating injury and incapacitating injury		Threshold between incapacitating injury and fatal injury	
	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat
Threshold constants			−1.09**	−2.33	0.235***	3.845	0.984***	21.8
Crash characteristics								
Crash involving pedestrian	2.23***	21.9						
Crash involving bikers	1.97***	18.3						
Roadway characteristics								
Traffic Signal	−0.423***	−6.89						
Stop Sign	−0.954***	−5.11						
Curve and Grade	1.27***	5.409						
Wet road	0.293***	2.94						
Single occupancy vehicle hourly traffic flow	−0.0003***	−3.29						
High occupancy vehicle hourly traffic flow	0.0008***	2.91						
Truck hourly traffic flow	0.0011***	5.51						
Commercial van hourly traffic flow	−0.0007***	−2.96					0.0008***	4.15
Vehicle characteristics								
Mean vehicle weight	0.026***	10.7			0.011***	12.9	−0.0035***	−3.812
Environmental characteristics								
Dawn	−0.304**	−1.79						
Dusk	−0.531**	−2.83						
Snow	−0.493*	−2.46						
Rain					0.15***	2.11		
Built environment characteristics								
Service job density							−0.042**	−1.74
Entertainment job density							0.058**	1.77
Time of day effects								
PM	−0.389***	−2.75			−0.325***	−3.47		
Night time	−0.565***	−4.23			−0.472***	−5.46		
Spatial parameter	0.105***	−6.5						
Log-composite likelihood at convergence: −1808.1								
Number of observations	735							

***, **, * Significance at 1%, 5%, 10% level.

for both models are summarized in Tables 4 and 5. Based on the spatial auto-correlation analysis result and the available data, spatial generalized ordered probit models were further added, and the estimates are summarized in Tables 6 and 7.

4.2. Estimation results analysis

In order to tell the full story of the effect of a variable on the probability of an crash severity level, marginal effects of all the selected explanatory variables were estimated for the random parameter models. In all the tables, the base level for categorical variables is presented in parenthesis. For example, for the time of day of the crash, the base level is “AM (6:00 A.M.–10:00 A.M.)”. For the goodness of fit of the model, the log-likelihoods and AICs are estimated for the random parameter ordered probit models, and the composite log-likelihoods are estimated for the spatial generalized ordered probit models. The effects of variables from each category are discussed in detail in the following sections.

4.2.1. Vehicle characteristics: random effect of vehicle registered weight

The heterogeneous effect of truck registered weight was found to be significant in both single-vehicle and multi-vehicle crashes; however, the estimates from the two models are different, as shown in Tables 4 and 5. In multi-vehicle crashes, the coefficient of the truck weight is normally distributed with a mean of 0.0016 and a variance of 0.000124. The result demonstrates a positive effect of truck weight difference on multi-vehicle crash severity with relatively small variance (similar to (Kockelman and Kweon, 2002; Lemp et al., 2011)). In order words, it is unnecessary to consider the truck weight difference as a random parameter in multi-vehicle crashes. The effect of truck weight on crash severity is found to be positive (with a mean of 0.0086) in single-vehicle crashes, with a larger variance (0.01), indicating that the effect of truck weight on crash severity is more heterogeneous in single-vehicle crashes. The marginal estimates show that with a 1000 lb increase in truck weight difference, the probability of the crash involving non-incapacitating injury will decrease 0.227%, and the probability of involving fatality will increase 0.053%. One further reason to be very careful considered in concluding the heterogeneous effect of vehicle weight on truck crash severity is the fact that vehicle weight is probably correlated with many compounding variables. For example, drivers for heavier trucks may have more strict training than light truck drivers.

4.2.2. Crash characteristics: traffic control condition, weather condition, roadway design attributes, light condition and road surface condition

The presence of traffic signal shows similar effects on crash severity in both single-vehicle crashes and multi-vehicle crashes: the crashes tend to be more severe when traffic sign/signal or a stop sign was presented at the crash location, which may be explained by the aggressive driving behavior exhibited at traffic signals (e.g., drivers tend to speed up when they see the traffic light turning red), and the fact that locations with traffic sign/signals tend to have more complex traffic conditions (e.g., intersections). The road surface condition also generates similar effects on both single and multi-vehicle crashes: trucks driving on wet roads tend to have

Table 7
Spatial Generalized Ordered Probit Model Estimation Results (Multi-vehicle Crashes).

Variables	Latent injury risk propensity		Threshold between property damage only and non-incapacitating injury		Threshold between non-incapacitating injury and incapacitating injury		Threshold between incapacitating injury and fatal injury	
	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat
Threshold constants			−14.9	−0.337	−0.674***	−57.256	0.893***	59.432
Roadway characteristics								
Stop Sign	0.268***	9.919						
Curve and Grade	−0.386***	−10.99						
Wet road	0.222***	9.363						
Commercial van hourly traffic flow	−0.0003***	−5.346					0.0008***	4.15
Vehicle characteristics								
Truck weight difference between the maximum Truck weigh and the minimum truck weight	0.0026***	17.946			0.011***	12.9	−0.0035***	−3.812
Environmental characteristics								
Dark road with lights	−0.386***	−8.57						
Rain	−0.079***	−2.823						
Built environment characteristics								
Retail job density	0.0098***	6.312						
Time of day effects								
PM	0.095***	6.921			−0.021	0.926		
Night time	0.173***	7.560						
Spatial parameter	0.178***	−15.3						
Log-composite likelihood at convergence: −46,296								
Number of observations	3769							

***, **, * Significance at 1%, 5%, 10% level.

more severe injuries. Cautious driving behaviors under adverse weather conditions (e.g., snow or rain) tend to reduce the injury severity levels in both single and multi-vehicle crashes.

There are a number of variables influencing the crash severity in single or multi-vehicle crashes only. For example, single-vehicle crashes involving pedestrian/bikers are usually more serious, which could be explained by the fact that pedestrians and bikers are more vulnerable in car crashes. Single-vehicle crashes also show less severe injuries under bad light conditions (dawn/dusk), which may be explained by the cautious driving behaviors under bad lighting conditions. For multi-vehicle crashes, dark roads with lighting help to reduce the injury severities.

4.2.3. *Traffic volume effects*

Four out of the six traffic volume variables (single occupancy vehicle, high occupancy vehicle, truck and commercial van) were found significant in single-vehicle crashes. The adverse effects of traffic volume of the sov and commercial vans on crash severity are consistent with the existing literature, and can be explained as a consequence of cautious driving (Chen and Chen, 2011; Christoforou et al., 2010; Das et al., 2009). It is interesting to see that the increase in traffic volume for high occupancy vehicles may lead to more severe crashes, possibly as an outcome of more people involved in the crashes.

4.2.4. *Built environment effect*

For single-vehicle crashes, the job density for service employment affected the threshold between incapacitating injury and fatal injury negatively, and the job density for entertainment employment had the opposite effect. New York City has many entertainment industries located at multiple city centers (e.g., the Broadway district and Time Square), where the dense population (especially tourists) and traffic are encountered; therefore it is obvious to see that more severe crashes are more likely to occur in those locations. For multi-vehicle crashes, areas with high retail job densities tend to have more severe injuries, which could be explained by similar reasons behind the effect of entertainment job densities mentioned previously.

4.2.5. *Time of day effect*

The time of day indicator variables are found to be statistically significant for both single-vehicle and multi-vehicle crashes in the pm and night periods. Interesting, the effects are opposite in these two types of crashes: single-vehicle crash tends to less severe during the afternoon and night periods, while multi-vehicle crashes are more dangerous during these periods. The finding is meaningful for policy makers considering special time delivery windows in the big cities. For example, knowing that truck crashes occurring at night are not statistically more severe than those that occur during normal delivery hours, policy makers may encourage shippers (who originally may have safety concerns for late night deliveries) to shift their deliveries to off-hours, and thus relieve the traffic congestion problems during daytime and reduce the corresponding economic losses.

4.2.6. *Spatial dependency effect*

Tables 6 and 7 show the model estimates when the spatial dependency effect is taken into account in single-vehicle and multi-vehicle crashes. The results are consistent with the estimates from the random coefficient single-vehicle crash model. The spatial parameters are statistically significant in both models, indicating that individual crashes in truck crashes are indeed correlated with each other, however, the spatial dependency effect fades rapidly with distance, which is indicated by the inverse of the cubed distance weight matrix.

4.2.7. Limitation of the current study

Although every effort was made through the study to account for all possible influencing factors of truck crashes, a number of attributes such as collision type, belt use, and alcohol consumption, are not included due to the limited information provided by the New York State's DOT. It should also be acknowledged that this study focus on exploring the time of day effect on truck crashes, while it could be extended to more comprehensive continuous temporal effects in future studies.

5. Conclusions

The study used two random parameter ordered probit models to explore the influencing factors of single-vehicle and multi-vehicle crashes separately. Two spatial generalized ordered probit models also employed to account for the potential spatial dependency amongst single-vehicle and multi-vehicle crashes. The results assessed a wide range of influencing factors of truck crash severity, including temporal and spatial dependency, crash characteristics, roadway characteristics, built environment variables, driver and occupant characteristics, environmental factors and vehicle characteristics. The empirical results clearly reveal that single-vehicle truck crashes and multi-vehicle crashes have different influencing factors, and there are statistically significant heterogeneous effects of truck weights on crash severity for both models. The spatial dependency and the temporal effect are also found statistically significant in this context; such finding provides useful insights for transportation modelers and planners in terms of planning and making policy decisions.

Acknowledgement

The research reported in this paper was supported by the University Transportation Research Center Region II as part of the Project "Investigating Temporal Effects on Truck Crash Occurrence and Severity Level in New York City."

References

Öström, M., Eriksson, A., 1993. Single-vehicle crashes and alcohol: a retrospective study of passenger car fatalities in northern Sweden. Crash Anal. Prev. 25 (2), 171–176.
Abdel-Aty, M., Wang, X., 2006. Crash estimation at signalized intersections along corridors: analyzing spatial effect and identifying significant factors. Transp. Res. Rec.: J. Transp. Res. Board 1953, 98–111.
Aguero-Valverde, J., Jovanis, P.P., 2006. Spatial analysis of fatal and injury crashes in Pennsylvania. Crash Anal. Prev. 38 (3), 618–625.
Aguero-Valverde, J., Jovanis, P., 2008. Analysis of road crash frequency with spatial models. Transp. Res. Rec.: J. Transp. Res. Board 2061, 55–63.
Aguero-Valverde, J., Jovanis, P., 2010. Spatial correlation in multilevel crash frequency models: effects of different neighboring structures. Transp. Res. Rec.: J. Transp. Res. Board 2165, 21–32.
Aguero-Valverde, J., 2013. Multivariate spatial models of excess crash frequency at area level: case of Costa Rica. Crash Anal. Prev. 59, 365–373.
Amoros, E., Martin, J.L., Laumon, B., 2003. Comparison of road crashes incidence and severity between some French counties. Crash Anal. Prev. 35 (4), 537–547.
Aziz, H.A., Ukkusuri, S.V., Hasan, S., 2013. Exploring the determinants of pedestrian–vehicle crash severity in New York City. Crash Anal. Prev. 50, 1298–1309.
Barua, S., El-Basyouny, K., Islam, M.T., 2016. Multivariate random parameters collision count data models with spatial heterogeneity. Anal. Methods Crash Res. 9, 1–15.
Bedard, M., Guyatt, G.H., Stones, M.J., Hirdes, J.P., 2002. The independent contribution of driver, crash, and vehicle characteristics to driver fatalities. Crash Anal. Prev. 34 (6), 717–727.
Bhat, C.R., Born, K., Sidharthan, R., Bhat, P.C., 2014. A count data model with endogenous covariates: formulation and application to roadway crash frequency at intersections. Anal. Methods Crash Res. 1, 53–71.
Brinckerhoff, P., (2009). New York Best Practice Model (NYBPM) For Regional Travel Demand Forecasting.
Castro, M., Paleti, R., Bhat, C.R., 2012. A latent variable representation of count data models to accommodate spatial and temporal dependence: application to predicting crash frequency at intersections. Transp. Res. B: Methodol. 46 (1), 253–272.
Castro, M., Paleti, R., Bhat, C.R., 2013a. A spatial generalized ordered response model to examine highway crash injury severity. Accid. Anal. Prev. 52, 188–203, http://dx.doi.org/10.1016/j.aap.2012.12.009 (Research Support, Non-U.S. Gov't).
Castro, M., Paleti, R., Bhat, C.R., 2013b. A spatial generalized ordered response model to examine highway crash injury severity. Crash Anal. Prev. 52, 188–203, http://dx.doi.org/10.1016/j.aap.2012.12.009 (Research Support, Non-U.S. Gov't).
Chang, L.-Y., Chen, W.-C., 2005. Data mining of tree-based models to analyze freeway crash frequency. J. Saf. Res. 36 (4), 365–375.
Chen, F., Chen, S., 2011. Injury severities of truck drivers in single-and multi-vehicle crashes on rural highways. Crash Anal. Prev. 43 (5), 1677–1688, http://dx.doi.org/10.1016/j.aap.2011.03.026 (Research Support, U.S. Gov't, Non-P.H.S.).
Chiou, Y.-C., Fu, C., 2015. Modeling crash frequency and severity with spatiotemporal dependence. Anal. Methods Crash Res. 5, 43–58.
Christoforou, Z., Cohen, S., Karlaftis, M.G., 2010. Vehicle occupant injury severity on highways: an empirical investigation. Crash Anal. Prev. 42 (6), 1606–1620, http://dx.doi.org/10.1016/j.aap.2010.03.019.
Cressie, N., 2015. Statistics for Spatial Data. John Wiley & Sons.
Das, A., Abdel-Aty, M., Pande, A., 2009. Using conditional inference forests to identify the factors affecting crash severity on arterial corridors. J. Saf. Res. 40 (4), 317–327, http://dx.doi.org/10.1016/j.jsr.2009.05.003.
Dubin, R.A., 1988. Estimation of regression coefficients in the presence of spatially autocorrelated error terms. Rev. Econ. Stat., 466–474.
Dupont, E., Papadimitriou, E., Martensen, H., Yannis, G., 2013. Multilevel analysis in road safety research. Crash Anal. Prev. 60, 402–411.
El-Basyouny, K., Sayed, T., 2009. Urban arterial crash prediction models with spatial effects. Transp. Res. Rec.: J. Transp. Res. Board 2102, 27–33.
Eluru, N., Bhat, C.R., 2007. A joint econometric analysis of seat belt use and crash-related injury severity. Crash Anal. Prev. 39 (5), 1037–1049.
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. Crash Anal. Prev. 40 (3), 1033–1054, http://dx.doi.org/10.1016/j.aap.2007.11.010 (Research Support, Non-U.S. Gov't).
Eluru, N., Paleti, R., Pendyala, R., Bhat, C., 2010. Modeling injury severity of multiple occupants of vehicles: copula-based multivariate approach. Transp. Res. Rec.: J. Transp. Res. Board 2165, 1–11.
F.M.C.S. Administration, (2013). Commercial Motor Vehicle Facts–2013.
Federal Highway Administration, 2012. Federal Grant Opportunity Request for Applications (RFA): Off Hours Freight Delivery Pilot Project (Retrieved July 1, 2012, from http://apply07.grants.gov/apply/opportunities/instructions/oppDTFH61-12-RA-00016-cfda20.200-cidDTFH61-12-RA-00016-instructions.pdfhttp://gallery.mailchimp.com/f0a61517ea47425a65df2fff6/files/_450_000_Funding_Announcement_Application_Due_July_13_2012.pdf?utm_source=NYLCHVCC+Stakeholders&utm_campaign=a5f9271229-%24450%2C000_Funding_Annoucement&utm_medium=email).
Flask, T., Schneider, W., 2013. A Bayesian analysis of multi-level spatial correlation in single vehicle motorcycle crashes in Ohio. Saf. Sci. 53, 1–10.
Gandhi, U., 2014. QGIS Tutorials and Tips (Retrieved from http://www.qgistutorials.com/en/docs/performing_spatial_joins.html).
Geedipally, S.R., Lord, D., 2010. Investigating the effect of modeling single-vehicle and multi-vehicle crashes separately on confidence intervals of Poisson–gamma models. Crash Anal. Prev. 42 (4), 1273–1282.
Golob, T.F., Recker, W.W., Leonard, J.D., 1987. An analysis of the severity and incident duration of truck-involved freeway crashes. Crash Anal. Prev. 19 (5), 375–395.
Griffith, M., 1999. Safety evaluation of rolled-in continuous shoulder rumble strips installed on freeways. Transp. Res. Rec.: J. Transp. Res. Board 1665, 28–34.
Guo, F., Wang, X., Abdel-Aty, M.A., 2010. Modeling signalized intersection safety with corridor-level spatial correlations. Crash Anal. Prev. 42 (1), 84–92.
Hakim, D., 2004. Average US Car Is Tipping Scales at 4000 Pounds. NY Times (May, 5).
Holguín-Veras, J., Ozbay, K., Kornhauser, A., Shorris, A., Ukkusuri, S., 2010a. Integrative Freight Demand Management in the New York City Metropolitan Area. Center for Infrastructure, Transportation and the Environment, Rensselaer Polytechnic Institute, Troy, NY.
Holguín-Veras, J., Ozbay, K., Kornhauser, A.L., Shorris, A., Ukkusuri, S., (2010). Integrative Freight Demand Management in the New York City Metropolitan Area.
Islam, M., Hernandez, S., 2013a. Large truck–Involved crashes: exploratory injury severity analysis. J. Transp. Eng. 139 (6), 596–604.
Islam, M., Hernandez, S., 2013b. 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., 2013c. Multi-vehicle collisions involving large trucks on highways: an exploratory discrete outcome analysis. J. Transp. Eng 139, 596–604.
Ivan, J.N., Pasupathy, R.K., Ossenbruggen, P.J., 1999. Differences in causality factors for single and multi-vehicle crashes on two-lane roads. Crash Anal. Prev. 31 (6), 695–704.
Ivan, J.N., Wang, C., Bernardo, N.R., 2000. Explaining two-lane highway crash rates using land use and hourly exposure. Crash Anal. Prev. 32 (6), 787–795.
Ivan, J., 2004. New approach for including traffic volumes in crash rate analysis and forecasting. Transp. Res. Rec.: J. Transp. Res. Board 1897, 134–141.
Jovanis, P.P., Chang, H.-L., 1986. Modeling the relationship of crashes to miles traveled. Transp. Res. Rec. 1068, 42–51.

Kahane, C.J., (2003). Vehicle weight, fatality risk and crash compatibility of model year 1991–99 passenger cars and light trucks.

Karlaftis, M.G., Tarko, A.P., 1998. Heterogeneity considerations in crash modeling. Crash Anal. Prev. 30 (4), 425–433.

Khattak, A.J., Schneider, R.J., Targa, F., 2003. Risk factors in large truck rollovers and injury severity: analysis of single-vehicle collisions. Presented at 2003 Annual Transportation Research Board Meeting, on TRB CD-ROM.

Kim, J.-K., Ulfarsson, G.F., Kim, S., Shankar, V.N., 2013. Driver-injury severity in single-vehicle crashes in California: a mixed logit analysis of heterogeneity due to age and gender. Crash Anal. Prev. 50, 1073–1081.

Kockelman, K.M., Kweon, Y.-J., 2002. Driver injury severity: an application of ordered probit models. Crash Anal. Prev. 34 (3), 313–321 (Research Support, Non-U.S. Gov't Review).

LeSage, J.P., Pace, R.K., 2009. Introduction to Spatial Econometrics. Boca Raton, FL, CRC Press, Taylor & Francis Group.

LeSage, J.P., 2009. Spatial probit models, estimation and interpretation. In: Paper Presented at the 3rd World Conference of the Spatial Econometrics Association, Barcelona, Spain (keynote address retrieved from).

Lee, C., Abdel-Aty, M., 2005. Comprehensive analysis of vehicle–pedestrian crashes at intersections in Florida. Crash Anal. Prev. 37 (4), 775–786.

Lemp, J.D., Kockelman, K.M., Unnikrishnan, A., 2011. Analysis of large truck crash severity using heteroskedastic ordered probit models. Crash Anal. Prev. 43 (1), 370–380, http://dx.doi.org/10.1016/j.aap.2010.09.006.

Leslie, T.F., Kronenfeld, B.J., 2011. The colocation quotient: a new measure of spatial association between categorical subsets of points. Geogr. Anal. 43 (3), 306–326.

Lord, D., Mannering, F., 2010. The statistical analysis of crash frequency data: a review and assessment of methodological alternatives. Transp. Res. A: Policy Pract. 44, 297–305.

Lord, D., Persaud, B.N., 2000. Crash prediction models with and without trend – Application of the generalized estimating equations procedure. Highw. Traffic Saf.: Crash Data Anal. Tools Stat. Methods 1717, 102–108.

Lord, D., Washington, S.P., Ivan, J.N., 2005. Poisson, Poisson-gamma and zero-inflated regression models of motor vehicle crashes: balancing statistical fit and theory. Crash Anal. Prev. 37 (1), 35–46.

Lui, K.-J., McGee, D., Rhodes, P., Pollock, D., 1989. An application of a conditional logistic regression to study the effects of safety belts, principal impact points, and car weights on drivers' fatalities. J. Saf. Res. 19 (4), 197–203.

MacNab, Y.C., 2004. Bayesian spatial and ecological models for small-area crash and injury analysis. Crash Anal. Prev. 36 (6), 1019–1028.

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

Martensen, H., Dupont, E., 2013. Comparing single vehicle and multivehicle fatal road crashes: a joint analysis of road conditions: time variables and driver characteristics. Crash Anal. Prev. 60, 466–471.

Mensah, A., Hauer, E., 1998. Two problems of averaging arising in the estimation of the relationship between crashes and traffic flow. Transp. Res. Rec.: J. Transp. Res. Board 1635, 37–43.

Miller, T., Levy, D., Spicer, R., Lestina, D., 1998. Allocating the costs of motor vehicle crashes between vehicle types. Transp. Res. Rec.: J. Transp. Res. Board 1635, 81–87.

Miller, T.R., Spicer, R.S., Lestina, D.C., Levy, D.T., 1999. Safest to travel by bicycle, car, or big truck? Traffic Inj. Prev. 1 (1), 25–34.

Milton, J.C., Shankar, V.N., Mannering, F.L., 2008. Highway crash severities and the mixed logit model: an exploratory empirical analysis. Crash Anal. Prev. 40 (1), 260–266.

Mitra, S., 2009. Spatial autocorrelation and Bayesian spatial statistical method for analyzing intersections prone to injury crashes. Transp. Res. Rec.: J. Transp. Res. Board 2136, 92–100.

Morgan, A., Mannering, F.L., 2011. The effects of road-surface conditions, age, and gender on driver-injury severities. Crash Anal. Prev. 43 (5), 1852–1863.

NYMTC, 2013. The Basics of Freight Transportation in the New York Region. The New York Metropolitan Transportation Council, New York, New York.

NYSDOT, N. Y. S. D. o. T. (Producer). (2014). New York Sate Department of Motor Vehicles Retrieved from: http://dmv.ny.gov/about-dmv/statistical-summaries.

Ni, L., Wang, X.C., Zhang, D., 2016. Impacts of information technology and urbanization on less-than-truckload freight flows in China: an analysis considering spatial effects. Transp. Res. Part A Policy Pract. 92, 12–25.

Noland, R.B., Quddus, M.A., 2004. A spatially disaggregate analysis of road casualties in England. Crash Anal. Prev. 36 (6), 973–984.

Pahukula, J., Hernandez, S., Unnikrishnan, A., 2015. A time of day analysis of crashes involving large trucks in urban areas. Crash Anal. Prev. 75, 155–163.

Paleti, R., Eluru, N., Bhat, C.R., 2010. Examining the influence of aggressive driving behavior on driver injury severity in traffic crashes. Crash Anal. Prev. 42 (6), 1839–1854.

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. Crash Anal. Prev. 38 (6), 1071–1080.

Quddus, M.A., 2008. Modelling area-wide count outcomes with spatial correlation and heterogeneity: an analysis of London crash data. Crash Anal. Prev. 40 (4), 1486–1497.

Sarrias, M., (2014). Discrete Choice Models with Random Parameters in R: The Rchoice Package.

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. Crash Anal. Prev. 43 (5), 1666–1676, http://dx.doi.org/10.1016/j.aap.2011.03.025 (Review).

Schrank, D., Eisele, B., Lomax, T., 2012. 2012 Urban Mobility Report. Texas A&M R Transportation Institute, College Station, Texas.

Shankar, V., Mannering, F., Barfield, W., 1995. Effect of roadway geometrics and environmental factors on rural freeway crash frequencies. Crash Anal. Prev. 27 (3), 371–389.

US EPA, (Producer) (2014). Smart Location Database. Retrieved from: http://www.epa.gov/smartgrowth/smartlocationdatabase.htm.

Venkataraman, N., Ulfarsson, G., Shankar, V., Oh, J., Park, M., 2011. Model of relationship between interstate crash occurrence and geometrics: exploratory insights from random parameter negative binomial approach. Transp. Res. Rec.: J. Transp. Res. Board 2236, 41–48.

Wang, X., Kockelman, K., 2005a. Use of heteroscedastic ordered logit model to study severity of occupant injury: distinguishing effects of vehicle weight and type. Transp. Res. Rec.: J. Transp. Res. Board 1908, 195–204.

Wang, X., Kockelman, K.M., 2005b. Use of heteroscedastic ordered logit model to study severity of occupant injury: distinguishing effects of vehicle weight and type. Transp. Res. Rec.: J. Transp. Res. Board 1908 (1), 195–204.

Wang, C., Quddus, M.A., Ryley, T., Enoch, M., Davison, L., 2012. Spatial models in transport: a review and assessment of methodological issues. Paper Presented at the Transportation Research Board 91 St Annual Meeting.

White, M.J., 2003. The Arms Race on American Roads: The Effect of SUV's and Pickup Trucks on Traffic Safety. NBER Working Paper.

Wu, Q., Chen, F., Zhang, G., Liu, X.C., Wang, H., Bogus, S.M., 2014. Mixed logit model-based driver injury severity investigations in single-and multi-vehicle crashes on rural two-lane highways. Crash Anal. Prev. 72, 105–115.

Yasmin, S., Eluru, N., 2013. Evaluating alternate discrete outcome frameworks for modeling crash injury severity. Crash Anal. Prev. 59, 506–521.

Yasmin, S., Eluru, N., Bhat, C.R., Tay, R., 2014. A latent segmentation based generalized ordered logit model to examine factors influencing driver injury severity. Anal. Methods Crash Res. 1, 23–38.

Yu, R., Abdel-Aty, M., 2013. Multi-level Bayesian analyses for single-and multi-vehicle freeway crashes. Crash Anal. Prev. 58, 97–105.

Zhang, D., Wang, X.C., 2014. Transit ridership estimation with network Kriging: a case study of Second Avenue Subway, NYC. J. Transp. Geogr. 41, 107–115.

Zhang, D., Wang, X., 2015a. Investigating the dynamic spillover effects of low-cost airlines on airport airfare through spatio-temporal regression models. Networks Spatial Econ., 1–16.

Zhang, D., Wang, X., 2015b. Effects of jobs-residence balance on commuting patterns: differences in employment sectors and urban forms. Transp. Res. Rec: J. Transp. Res. Board 2500, 80–92.

Zhang, D., Wang, X.C., 2016. Analyses considering partner selection and joint decision making: investigation of freight demand with spatial matching models. In Transportation Research Board 95th Annual Meeting, 16–6449.

Zhang, D., Magalhães, D.J.A.V., Wang, X.C., 2014. Prioritizing bicycle paths in Belo Horizonte City, Brazil: analysis based on user preferences and willingness considering individual heterogeneity. Transp. Res. Part A Policy Pract. 67, 268–278.

Zhu, X., Srinivasan, S., 2011. A comprehensive analysis of factors influencing the injury severity of large-truck crashes. Crash Anal. Prev. 43 (1), 49–57.

Zou, Y., Zhang, Y., Lord, D., 2013. Application of finite mixture of negative binomial regression models with varying weight parameters for vehicle crash data analysis. Crash Anal. Prev. 50, 1042–1051.

Zou, Wei, Wang, Xiaokun, Zhang, Dapeng, 2015. Truck accident severity in new York city: an investigation of the spatiotemporal effects and vehicle weight. Transportation Research Board 94th Annual Meeting, No.