Final Project Report

Addressing potentially missing relevant information on attitudes and other behavioral elements as unobserved heterogeneity in highway safety studies

Prepared for **Teaching Old Models New Tricks (TOMNET) Transportation Center**

By,

Fred Mannering
Email: flm@usf.edu

Asim Alogaili
Email: asimalogaili@usf.edu

Suryaprasanna Kumar Balusu
Email: balusu@mail.usf.edu

Department of Civil and Environmental Engineering
University of South Florida
4202 E. Fowler Avenue, Tampa, FL 33620

June 2022

TECHNICAL REPORT DOCUMENTATION PAGE

<table>
<tr><td>1. Report No.
N/A</td><td>2. Government Accession No.
N/A</td><td colspan="2">3. Recipient's Catalog No.
N/A</td></tr>
<tr><td colspan="2" rowspan="2">4. Title and Subtitle
Addressing potentially missing relevant information on attitudes and other behavioral elements as unobserved heterogeneity in highway safety studies</td><td colspan="2">5. Report Date
June 2022</td></tr>
<tr><td colspan="2">6. Performing Organization Code
N/A</td></tr>
<tr><td colspan="2">7. Author(s)
Fred Mannering, https://orcid.org/0000-0002-2803-4582
Asim Alogaili, https://orcid.org/0000-0002-8268-7461
Suryaprasanna K. Balusu, https://orcid.org/0000-0003-0726-4036</td><td colspan="2">8. Performing Organization Report No.
N/A</td></tr>
<tr><td colspan="2" rowspan="2">9. Performing Organization Name and Address
University of South Florida, 4202 E. Fowler Avenue, Tampa, FL 33620</td><td colspan="2">10. Work Unit No. (TRAIS)
N/A</td></tr>
<tr><td colspan="2">11. Contract or Grant No.
69A3551747116</td></tr>
<tr><td colspan="2" rowspan="2">12. Sponsoring Agency Name and Address
U.S. Department of Transportation,
University Transportation Centers Program,
1200 New Jersey Ave, SE, Washington, DC 20590</td><td colspan="2">13. Type of Report and Period Covered
Research Report
(September 2020 – August 2021)</td></tr>
<tr><td colspan="2">14. Sponsoring Agency Code
USDOT OST-R</td></tr>
<tr><td colspan="4">15. Supplementary Notes
N/A</td></tr>
<tr><td colspan="4">16. Abstract
Understanding highway safety requires an assessment of physical and behavioral factors that influence the occurrence and resulting injury severity of crashes. This assessment is exceedingly difficult because it involves an understanding of wide variety of factors including human responses to external stimuli and complex interactions between the vehicle, roadway features, roadway conditions, traffic-related factors, and environmental conditions. In addition, attitudes and other factors that potentially affect drivers' safety-risk profiles play a key role in the occurrence and resulting injury severity of crashes. Recognizing that it would be virtually impossible to collect all of the data that could potentially influence the occurrence and severity of crashes, the safety field has moved forward by addressing omitted data as unobserved heterogeneity – using mixing distributions to arrive at appropriate statistical estimations. This study explores mixing distribution approaches to gain insight on the potential role that attitudes and other behavioral elements may play. This is done by analyzing pedestrian injuries over time (using data from Kansas) and assessing the frequency and severity of crashes using data from roadway segments in multiple states. The findings indicate that mixing distributions provide an approach that can capture important attitudinal and other behavioral elements when such data cannot be realistically collected, and that mixing distributions should be used in highway-safety practice as quickly as possible through their use in the American Association of State Highway and Transportation Official's Highway Safety Manual.</td></tr>
<tr><td colspan="2">17. Key Words
Pedestrian injury, Random parameters logit, Temporal instability, Out-of-sample prediction, Unobserved heterogeneity, Spatial effects, Bivariate negative binomial model</td><td colspan="2">18. Distribution Statement
No restrictions.</td></tr>
<tr><td>19. Security Classif.(of this report)
Unclassified</td><td>20. Security Classif.(of this page)
Unclassified</td><td>21. No. of Pages
96</td><td>22. Price
N/A</td></tr>
</table>

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

DISCLAIMER

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated in the interest of information exchange. The report is funded, partially or entirely, by a grant from the U.S. Department of Transportation's University Transportation Centers Program. However, the U.S. Government assumes no liability for the contents or use thereof.

ACKNOWLEDGMENTS

This study was funded by a grant from A USDOT Tier 1 University Transportation Center, supported by USDOT through the University Transportation Centers program. The authors would like to thank the TOMNET, USDOT, and University of South Florida for their support of university-based research in transportation, and especially for the funding provided in support of this project.

TABLE OF CONTENTS

List of Tables

List of Figures

EXECUTIVE SUMMARY

Part 1 of this study explores the differences between day and night pedestrian-injury severities in vehicle-pedestrian crashes over a five-year period using data from the state of Kansas. To account for the effect of possible missing attitudinal data (generally not present in traditional data sources), mixing distributions are introduced in statistical model estimation. Using this approach, separate statistical models (random parameters logit models with possible heterogeneity in the means and variances of the random parameters) were estimated for day and night crashes to examine different pedestrian injury severity outcomes (no visible injury, moderate injury, and severe injury). Likelihood ratio tests were conducted to explore the temporal stability of the model estimations over different times of day and years. Many variables affecting injury severities were considered in model estimation including time and location of accidents, in addition to information on environmental, roadway, crash, vehicle, driver, and pedestrian characteristics. The findings indicate that the factors affecting pedestrian injury severities did change over time but that there is a clear day-night difference in the resulting injury severities of pedestrians, with nighttime crashes consistently resulting in more severe injuries overtime. This suggests policies and technologies that seek to essentially replicate daytime conditions (improved illumination, infrared pedestrian detection in vehicles, etc.) in nighttime conditions could have considerable safety benefits. Using the estimated random parameters models, extensive out-of-sample prediction simulations are used to provide estimates of the potential benefits of such nighttime mitigation policies and technologies, as well as how daytime/nighttime pedestrian injury severity probabilities have been changing over time. Because of the statistical significance of random parameters in the model estimations, attitudes and possible behavioral shifts are likely playing key roles.

Part 2 explores a limitation of the in safety practice where attitudes and other behavioral elements are generally not considered. Again, viewing these elements as unobserved heterogeneity, random parameters bivariate models are estimated where different crash frequencies on the same roadway section can be modeled simultaneously provides the flexibility to consider the correlations between different injury severity crash frequencies. Specifically, a bivariate model and two separate univariate models were estimated for noninjury and injury crashes on freeways. The performances of both model structures were compared using various metrics. Univariate models had better performance for no-injury crash frequencies and bivariate model had better performance for injury crashes. Most importantly, both univariate and bivariate safety had statistically significant random parameters, suggesting that unobserved heterogeneity is playing an important role. Again, the performance of both random parameters bivariate models and random parameters univariate models suggest that their ability to capture missing attitudinal and behavioral elements through mixing-distributions is an important practical consideration in model estimation. Because attitudinal data is nearly impossible to gather with traditional crash-data sources, the importance of considering unobserved heterogeneity via mixing distributions or other methods is an empirical necessity.

The findings of this study show the urgent need to consider unobserved heterogeneity in highway safety practice as a way of capturing attitudes, behavioral elements and other factors that cannot be captured from traditional crash data sources.

PART I

Differences between day and night pedestrian-injury severities: Accounting for temporal and unobserved effects in prediction

1.1 Introduction

Worldwide, roadway mortality rates have increased over the recent years reaching over 1.3 million fatalities each year, with the vulnerable roadway users (pedestrians, cyclists, and motorcyclists) accounting for nearly half of the fatalities (WHO, 2018). In the United States, the number of pedestrian fatalities increased by over 50% in 2019 compared to 2009 while other-roadway fatalities increased by less than 1%, for the same period (FARS, 2019). The increase in the fatality numbers clearly shows that pedestrians are facing a greater risk of fatal injury in recent years, and many factors potentially play a role in this. For example, vehicle manufacturing designs (moving away from to less aerodynamic frontal areas which are likely to be less safe for pedestrians) and vehicle choices (the movement away from traditional passenger cars to larger sport utility vehicles and pickup truck) have changed over the years which may cause more severe outcomes for pedestrians hit by vehicles. In addition, pedestrians may be adversely impacted by many new safety features because drivers might feel safer and drive more aggressively thus imposing an adverse externality on pedestrian safety. This risk-compensating effect of safety features has been supported by past empirical research (Winston et. al., 2006).

The increase in pedestrian fatalities could also be associated with other factors. For example, the constant lighting-condition changes throughout different times of the day could affect the risk perception ability for both drivers and pedestrians and these perceptions could be evolving negatively over time. In the United States, pedestrian fatalities increased by 67% during nighttime and by 16% during daytime in the period from 2009 to 2019 (GHSA, 2020; FARS, 2019). Both drivers and pedestrians could face significant challenges relating to visibility at night, with drivers not appropriately adjusting their speed to compensate for reduce nighttime visibility, and pedestrians overestimating drivers' ability to see them at night (Tyrrell et al., 2004). The combination of these two factors may be changing over time in fundamental ways, perhaps influenced over time by increased use of cell phones and other distractions, or simply behavioral evolutions over time.

Improved nighttime lighting is one possible way to address the issue of nighttime pedestrian injury severities, but streetlights may have the adverse effect increasing some road-side accidents (hitting light poles) and they have obvious cost (both capital costs and maintenance) and energy-consumption disadvantages. As a consequence, in many locations drivers are relying on vehicle headlights alone to drive during dark hours. The problem with this is that research has found that drivers usually overdrive their headlamps, meaning that they drive faster than their headlights actually allow for them to notice pedestrian and undertake proper braking and evasive action (Green, 2020). In addition, general increases in speed limits in recent years may also be playing a role since increased speed tends to lower the vision-perception ability of drivers as their focus narrows down to the center of the road instead of being attentive to the approaching pedestrians on the sides of the road (Green, 2020). Regarding day/night differences, drivers' vision-perception is known to change during different times of day as well (Behnood and Mannering, 2019; Song et al., 2021). Therefore, studying and quantifying the differences, over time, between the severity of pedestrian accidents in day versus night conditions is important for developing effective injury-mitigation policies.

Using five-year single-pedestrian single-vehicle crash data, the goal of the current report is to explore the reasons behind the differences in pedestrian injuries in day and night conditions, and to explore the potential temporal instability of these factors over time. The current report is distinguished from previous work that has addressed the temporal instability of pedestrian injury severities in two important areas. First, the current report will explicitly consider day vs. night models, making the case (theoretically and empirically) that these two time periods should be considered separately in model estimation. Second, and perhaps more importantly, the current report will use out-of-sample prediction to determine the aggregate differences between day and night conditions as well the aggregate differences over time. Thus, in contrast to past work that has looked at how the marginal effects of explanatory variables have changed over time, this report will not only look at changing marginal effects but what the aggregate effect of all variables has on injury severities over time and between day and night conditions.

The following section of this part of the report provides a summary of previous research work that has addressed the association between time of day and pedestrian injury severity. Next, the data and methodology used in the study are presented and a set of temporal instability tests are conducted and discussed, and the model estimation results are then presented. Finally, the last two sections of this part of the report provide a prediction comparison of results (using an out-of-sample simulation approach) and study conclusions.

1.2 Literature review

The effects of time of day on accident-injury severity outcomes have been widely discussed in the literature (Plainis and Murray, 2002; Hao et al., 2016; Jägerbrand and Sjöbergh, 2016; Behnood and Mannering, 2019). In general, different times of the day have been found to have different effects on the injury severity outcomes, and a number of explanations for this finding have been put forward in the literature. For example, drivers typically have been found to be more likely fatigued during night hours and that results in drivers reacting more slowly, being less attentive to the surroundings, and having weakened decision-making skills. Also, people in general were found to make more accurate decisions earlier in the day while tending to make less accurate decisions later in the day (Leone et al., 2017). These findings are supported by the results of many previous accident-analysis studies. For example, daylight conditions were found to significantly decrease severe-injury crash outcomes in the mornings while dark conditions were found to increase severe-injury crash outcomes for specific years (Behnood and Mannering, 2019). It was also found that crashes occurring at night were associated with more severe-injury accidents compared to other times of the day (Hao et al., 2016). Furthermore, some work has found the time of the day and weather conditions to have an interactive effect on accidents injury severity with some factors found be consistent across different time of day and weather combinations (such as high speed roadways, tree-related accidents, and pedestrian-involved accidents) while other factors were found to be varying across combinations (such as vehicle type, vehicle age, driver age and gender, and surface condition) (Ariannezhad and Wu, 2019; Fountas et al., 2020).

Time of day effects in vehicle/pedestrian collisions have been recognized previously in the extant literature (Sullivan, 2001; Kim et al., 2010; Mokhtarimousavi, 2019; Mokhtarimousavi et al., 2020; Pantangi et al., 2021). The effects of time of day on pedestrian injury severity is seen to have more notable trends due to the high vulnerability of pedestrians to injury relative to other roadway users, and perceptual changes of pedestrian and driver behaviors between day and night conditions. Daylight indicator variables, for example, have been previously found to be associated with a decrease in the likelihood of severe pedestrian injuries and a corresponding increase in the likelihood of minor pedestrian injuries (Behnood and Mannering, 2016). Dark (both with and without roadway light), dusk, and dawn condition indicators have been found to be associated with higher probabilities of severe pedestrian injuries (Song et al., 2020), and pedestrian-involved accidents have been found to increase the likelihood of severe injury probability in most time of day and whether variations except for dark (lighted) roadways and fine or adverse weather (Fountas et al., 2020).

In general, non-daylight condition indicators have consistently been found to be associated with an increase in the likelihood of severe pedestrian injuries in pedestrian-involved vehicular accidents (Kim et al., 2010; Aziz et al., 2013; Pour-Rouholamin and Zhou, 2016; Xin et al., 2017; Chen and Fan, 2019; Mokhtarimousavi et al., 2020). This observed increase in the likelihood of severe injuries during non-daylight conditions might be due to several reasons including poor vision, higher speeds (due to potentially lower levels of congestion during non-day periods), driver-pedestrian fatigue, and walking or driving under influence (Kim et al., 2010; Aziz et al., 2013; Abay, 2013; Mohamed, 2013; Yasmin et al., 2014). Other reasons might be related to fundamental diurnal behavioral variations among both drivers and pedestrians. For example, drivers might accelerate or decelerate in higher rates because of the traffic patterns during certain times of the day (dawn and dusk) (Pantangi et al., 2021). Also, unexpected pedestrian speed changes may significantly affect the severity of vehicle-pedestrian accidents (Alhajyaseen and Iryo-Asano, 2017). At-fault pedestrians have been found to be associated with more severe injuries in vehicle-pedestrian crashes compared to at-fault drivers at both signalized and unsignalized intersections (Haleem et al., 2015). Another reason that could be associated with the increase in pedestrian injury severity at night could be the simple limitations of the illumination provided by vehicle headlight (Sullivan and Flannagan, 2011; Green, 2018).

There is a general consensus that better lighting conditions (thus more closely replicating daytime conditions) may reduce injury risks in vehicle-pedestrian accidents (Fountas et al., 2020) and that poor natural lighting can be enhanced with Li et al. (2021) finding that sufficient artificial lighting is safer for pedestrians and drivers relative to poor natural lighting conditions (dawn and dusk).

Other factors that have been found to be associated with higher likelihoods of severe pedestrian injuries are the day of the week (Pour-Rouholamin and Zhou, 2016; Li et al., 2021), older pedestrians (Yasmin et al., 2014; Haleem et al., 2015; Pour-Rouholamin and Zhou, 2016; Zamani et al., 2021), older drivers (Mohamed, 2013; Pour-Rouholamin and Zhou, 2016), younger drivers (Abay, 2013; Pour-Rouholamin and Zhou, 2016), male drivers (Forbes and Habib, 2015), male pedestrians (Abay, 2013), higher speed limits speed limit (Tay et al., 2011; Quin and Fan, 2021), multiple-lane road (Aziz et al., 2013; Pour-Rouholamin and Zhou, 2016), vehicle type (sport-utility vehicles and pickup) (Mohamed, 2013; Pour-Rouholamin and Zhou, 2016) and rainy weather (Tay et al., 2011; Aziz et al., 2013; Zhai et al., 2019). Some factors associated with a lower likelihood of severe pedestrian injury are the presence traffic signs, traffic signals, younger pedestrian (Zamani et al., 2021), pedestrian use of contrasting clothing (Pour-Rouholamin and Zhou, 2016), intersection-located accidents (Aziz et al., 2013; Mohamed, 2013), rush hours, and clear weather (Forbes and Habib, 2015). For an additional extensive review of factors affecting pedestrian injury severity in numerous previous studies, please see Zamani et al. (2021) as they cover a wide range of variables and crash-related attributes.

Possible temporal instability in the presence the unobserved heterogeneity in the statistical modeling of pedestrian injury severity outcomes in relation to day versus night crashes is a potentially important issue that has not yet been addressed in the literature. However, some previous work has discussed the changing effects of pedestrian-related accident injury severity outcome over time. Studying the variables that affect pedestrian injury severity and their temporal stability over a six-year period, Zamani et al. (2021) found that indicator variables affecting pedestrian injury severity were unstable over the years. They concluded that the daylight condition indicator was a temporally stable variable decreasing the likelihood of severe injury while the pedestrian injury severity during other lighting conditions (dark, dusk and dawn) were temporally unstable throughout the years. Behnood and Mannering (2016) found the effect of factors influencing the pedestrian injury severity to be unstable over multiple years using two different statistical methods (Behnood and Mannering, 2016). In their work, the daylight indicator variable was found to be mostly statistically significant over the eight-year period of their study. The behavior of this variable, however, was found to be unstable in regard to the marginal effects magnitudes. The reason for these observed temporal shifts in there is not clear, but the authors speculated that it may be associated with long-term underlying fundamental behavioral changes (Behnood and Mannering, 2016). It can also be associated specifically with changes in driver's behavior due to the effects of new vehicle safety technologies (with drivers offsetting these new safety features by driving more aggressively) and, therefore, this might have a great impact on the resulting pedestrian's injury severity due to their high vulnerability (Winston et al., 2006).

Batouli et al. (2020) also report some notable temporal changes in severe pedestrian injuries over an 11-year period. They found that the sport utility vehicle crash involvement was an increasing trend throughout the years with more than 20% difference, and they also found that pedestrian-impairment had an increasing effect on severe-injury outcomes over most of the years.

In general, changes in accident characteristics from year to year may introduce temporal shifts and that could be associated with several reasons like changes in decision making, risk taking, and cognitive behavior and reasoning biases. Ignoring such temporal changes may lead to inaccurate results and eventually ineffective crash-injury mitigation policies (Mannering, 2018).

Unobserved heterogeneity has become increasingly recognized as an issue likely to arise when using current traditional datasets to perform statistical analyses due to missing information (for example, detailed information on the speed and energy transfer through the vehicle, biomechanics at the time of collision, etc.) and the complex interactions among currently available information (Mannering et al., 2016). In previous studies, unobserved heterogeneity in injury-severity models has been accounted for using a wide variety of methodological approaches including mixed logit model, latent-class models, Markov-switching models, and random parameters logit models (Savolainen et al., 2011; Bhat and Mannering, 2014; Mannering et al., 2016; Alnawmasi and Mannering 2019; Alogaili and Mannering, 2020; Islam et al., 2020; Intini et al., 2020). Failing to account for unobserved heterogeneity might lead to inaccuracy when estimating and interpreting the results (Mannering et al., 2016; Washington et al., 2020). To get a sense of the range of methodological approaches that have been applied in pedestrian injury severity research in the past, and how unobserved heterogeneity was addressed, Zamani et al. (2021) provide an excellent summary of the methodological approaches used in recent pedestrian-injury severity research.

1.3 Empirical setting

The data used in this study are the pedestrian injury severities of crashes collected from Kansas over a five-year period from January 1, 2013, to December 31, 2017. The utilized data has a total of 1,849 single-pedestrian single-vehicle accidents with detailed information on the time and location of the crash, pedestrian and driver attributes (such as age, gender, injury severity, impairment condition, and action at the time of accident), vehicle characteristics (such as type, make, model, color, and year of manufacture), crash attributes (such as reason, type, direction, and place of the accident), and road and environment characteristics (such as speed limit, lighting condition, and pavement condition). The analysis herein focuses on estimating the models using the single-pedestrian single-vehicle crashes that occurred during two times of day, daytime (daylight lighting condition) and nighttime (dark, dawn and dusk lighting condition).

For the forthcoming statistical analysis, three possible injury-severity outcomes are considered: no visible injury (possible injury or no injury), moderate injury (non-incapacitating injury), and severe injury (fatality or incapacitating injury). For these injury levels, Table 1 provides the observed percent distribution of injuries by year and time of day (day vs. night) in the Kansas data. While the values in this table show the percent by injury level do fluctuate over time for both day and night accidents, there is not necessarily a clear increasing or decreasing trend. In contrast, in comparing day and night injury-level percentages across the years it is clear that nighttime accidents result in a consistently higher percentage of severe injuries relative to daytime accidents. The forthcoming statistical analysis will explore these yearly and daytime vs. nighttime trends in detail.

Table 1. Pedestrian injury severity distribution, percentages by year and time of day.

	Minor Injury		Moderate Injury		Severe Injury	
Year	Day	Night	Day	Night	Day	Night
2013	36.12	35.88	44.05	34.35	19.82	29.77
2014	46.29	42.86	41.92	34.78	11.79	22.36
2015	40.96	35.06	39.89	40.91	19.15	24.03
2016	42.42	25.83	43.56	40.40	14.02	33.12
2017	40.39	31.91	43.35	39.01	16.26	27.15

1.4 Methodological approach

In addition to the uncertainty of the interaction effects of pedestrian injury severity and day/night conditions, unobserved heterogeneity is expected to play a significant role in the study because factors that affect pedestrian injury severity are not fully covered in currently available data sets (Mannering et al., 2016). Thus, to undertake this issue, a statistical model that accounts for potentially complex forms of unobserved heterogeneity must be used to arrive at model estimates that are as accurate as possible. In this report, injury-severity probabilities are studied using a random parameters logit model that accounts for possible heterogeneity in the means and variances of random parameters, a textbook modeling approach that has become increasingly popular in recent empirical studies of injury severity (Washington et al., 2020).

For the case addressed in this report (the injury severity of pedestrians in single-vehicle single-pedestrian crashes) three possible injury-severity outcomes are considered: no visible injury (possible injury or no injury), moderate injury (non-incapacitating injury), and severe injury (fatality or incapacitating injury). Following past work (Washington et al., 2020), the modeling approach starts by defining a function that determines injury severity,

$$S_{kn} = \boldsymbol{\beta}_k \mathbf{X}_{kn} + \varepsilon_{kn} \tag{1}$$

where S_{kn} is an injury-severity function determining the probability of pedestrian-injury severity outcome k in vehicle-pedestrian crash n, $\mathbf{X}_{kn}$ is a vector of explanatory variables that affect pedestrian-injury severity level k, $\boldsymbol{\beta}_k$ is a vector of estimable parameters, and ε_{kn} is an error term. Assuming the error term is extreme value distributed, the standard multinomial logit model results as (McFadden, 1981),

$$P_n(k) = \frac{EXP(\boldsymbol{\beta}_k \mathbf{X}_{kn})}{\sum_{\forall k} EXP(\boldsymbol{\beta}_k \mathbf{X}_{kn})} \tag{2}$$

where $P_n(k)$ is the probability that crash n that will result in pedestrian-injury severity outcome k with K being the set of the three-possible injury-severity outcomes. To allow one or more parameter estimates in the vector $\boldsymbol{\beta}_k$ to vary across crash observations, Equation (2) can be rewritten, adding a mixing distribution, as (Train, 2009; Washington et al., 2020),

$$P_n(k) = \int \frac{EXP(\boldsymbol{\beta}_k \mathbf{X}_{kn})}{\sum_{\forall k} EXP(\boldsymbol{\beta}_k \mathbf{X}_{kn})} f(\boldsymbol{\beta}_k \mid \boldsymbol{\varphi}_k) d\boldsymbol{\beta}_k \tag{3}$$

Where $f(\boldsymbol{\beta}_k \mid \boldsymbol{\varphi}_k)$ is the density function of $\boldsymbol{\beta}_k$ and $\boldsymbol{\varphi}_k$ is a vector of parameters describing the density function (mean and variance), and all other terms are as previously defined.

The possibility of heterogeneity in the means and variances of random parameters is also considered by letting $\boldsymbol{\beta}_{kn}$ be a vector of estimable parameters that varies across crashes defined as (Mannering et al., 2016; Seraneeprakarn et al., 2017; Behnood and Mannering, 2017; Alnawmasi and Mannering, 2019; Behnood and Mannering, 2019; Islam et al., 2020; Washington et al., 2020; Al-Bdairi et al., 2020; Yu et al., 2020; Li et al., 2021; Hou et al., 2021; Yan et al., 2021; Song et al., 2021; Se et al., 2021; Zamani et al., 2021):

$$\boldsymbol{\beta}_{kn} = \beta + \boldsymbol{\Theta}_{kn}\mathbf{Z}_{kn} + \sigma_{kn} EXP(\omega_{kn}\mathbf{W}_{kn}) v_{kn} \tag{4}$$

where β is the mean parameter estimate across all crashes, $\mathbf{Z}_{kn}$ is a vector of crash-specific explanatory variables capturing heterogeneity in the mean that affects pedestrian injury-severity level k, $\boldsymbol{\Theta}_{kn}$ is the corresponding vector of estimable parameters, $\mathbf{W}_{kn}$ is the vector of crash-specific explanatory variables capturing heterogeneity in the standard deviation σ_{kn} with corresponding parameter vector ω_{kn}, and v_{kn} is a disturbance term. Possible correlation among random parameters was also considered (Fountas et al., 2018; Fountas et al., 2019; Saeed et al., 2019; Washington et al., 2020) during model estimation.

The models were estimated by simulated maximum likelihood with 1,000 Halton draws (McFadden and Train, 2000; Washington et al., 2020). To assist in the interpretation of the findings, marginal effects were also computed to capture the effect that a one-unit change in any specific explanatory variable has on the probability of an injury-severity outcome. The values of the corresponding marginal effects were calculated for each observation and were averaged over the population of observations.

1.5 Temporal stability tests

To statistically test if pedestrian-injury severity models are significantly different between day and night across the years from 2013 to 2017, likelihood ratio tests were conducted with the χ^2 distributed test statistic being (with degrees of freedom equal to the number of estimated parameters), for each year (see Washington et al., 2020),

$$X^2 = -2[LL(\boldsymbol{\beta}_{ND}) - LL(\boldsymbol{\beta}_D)] \tag{5}$$

where $LL(\boldsymbol{\beta}_{ND})$ is the log-likelihood at convergence of the model (for a given year) estimated using converged parameters from night data on day data (restricting the parameters to be night estimated parameters) for that year. $LL(\boldsymbol{\beta}_{D})$ is the log-likelihood at convergence of the model of the model using the day data. Reversing day and night for this test was also performed for each year, as well. The resulting value X^2 is χ^2 distributed and can be used to determine if the null hypothesis that the parameters are equal in the five years can be rejected. The results of the tests for day and night are shown in Table 2. The values in Table 2 show clearly that the null hypothesis that day and night injury severity models are the same can be rejected with over 99.9% for each of the five years.

Table 2. Likelihood ratio test results between day and night (in parantheses) for different years (refer to Equation 5).

Time	X^2	Percent Confidence Level	Degrees of Freedom
2013	98.399 (72.394)	99.99 (99.99)	17 (17)
2014	82.739 (260.759)	99.99 (99.99)	19 (16)
2015	45.911 (114.119)	99.99 (99.99)	11 (15)
2016	133.512 (72.120)	99.99 (99.99)	17 (11)
2017	351.043 (57.150)	99.99 (99.99)	12 (11)

Next, a series of likelihood ratio tests are conducted to determine if the separately estimated day models and separately estimated night models are temporally stable over the 5-year study period. The χ^2 distributed test statistic is now (with tests run separately for both day and night models),

$$X^2 = -2[LL(\boldsymbol{\beta}_{t_2 t_1}) - LL(\boldsymbol{\beta}_{t_1})] \tag{6}$$

where, $LL(\boldsymbol{\beta}_{t_2 t_1})$ is the log-likelihood at convergence of a model containing converged parameters based on using time-period t_2's data, while using data from time-period t_1, and

$LL(\beta_{t_1})$ is the log-likelihood at convergence of the model using time-period t_1's data (with parameters are no longer restricted to using time-period t_2's converged parameters as is the case for $LL(\beta_{t_2t_1})$. This test was also reversed such that time-period t_1's above becomes time period t_2 and time period t_2 above becomes subset t_1 (thus again giving two test results for each model comparison). The test results shown in Table 3 again indicate that the null hypothesis that day-model estimates are the same from one year to the next can be rejected with over 99% confidence and the same is true of the night-model estimates.[1]

The likelihood ratio test results in Tables 2 and 3 show clearly show that, to appropriately model the 5 years of available data, separate day and night models for each year should be estimated, for a total of 10 model (a day model for each of the 5 years and a night model for each of the 5 years).

1.6 Model estimation results

Below, we present a discussion of selected variables and their effects on pedestrian-injury severity in daytime and nighttime crashes over several years. Tables 4-13 present the estimation results for models estimated using 2013-2017 day and night pedestrian data, separately.[2] Also, Table 14 shows a comparison of marginal effects of single-pedestrian single-vehicle crashes over the four years for both times of day. Heterogeneity in the means of the random parameter was found to be significant in many estimations, but none of the models had statistically significant heterogeneity in the variance. Also, during estimation, models that

[1] An alternative likelihood ratio test, often included in the temporal instability literature, is to estimate a model using the data from all time periods and compare its log-likelihood at convergence to the log-likelihoods at convergence of models based on the individual time periods. As shown in Hou et al. (2022), this global stability test is less discriminating than the pairwise test shown in Equation 6, and it is more likely that overall instability will be found when there may be stability between year-pairs in the data. As a result of the Hou et al. (2022) findings and to save space, the global stability tests are not presented herein although all such tests were conducted and indicated overall global temporal instability.

[2] For determining the statistical significance of the standard deviations of random parameter it is important to note that the t-statistics are only suggestive, and the correct measure of statistical significance is the improvement in convergence as measured by the χ^2 distributed likelihood ratio test that compares the model with the parameter in question as fixed to a model where it is random model (where the additional standard deviation parameter is estimated). In all cases of the estimated models the null hypothesis that fixed and random parameters were equal could be rejected with over 90% confidence.

Table 3. Likelihood ratio test results between different years for day and night. Shown are the confidence levels in percentage and (degrees of freedom in parentheses) and [X^2 in brackets] (refer to Equation 6).

	Comparison year, t_2									
	2013		2014		2015		2016		2017	
Base year, t_1	Day	Night	Day	Night	Day	Night	Day	Night	Day	Night
2013	-	-	99.99 (19) [117.11]	99.99 (16) [471.35]	99.98 (11) [35.21]	99.99 (15) [67.64]	99.99 (16) [78.36]	99.99 (11) [46.97]	99.99 (16) [52.76]	99.99 (11) [69.12]
2014	99.99 (16) [87.85]	99.99 (17) [47.40]	-	-	99.92 (11) [31.90]	99.99 (15) [75.46]	99.99 (16) [51.41]	99.99 (11) [75.03]	99.99 (16) [284.28]	99.99 (11) [54.71]
2015	99.99 (16) [605.90]	99.99 (17) [74.37]	99.99 (19) [87.46]	99.99 (16) [86.54]	-	-	99.99 (16) [244.74]	99.99 (11) [79.42]	99.99 (16) [39.91]	99.99 (11) [46.77]
2016	99.99 (16) [71.35]	99.99 (17) [49.96]	99.99 (19) [62.90]	99.99 (16) [95.91]	99.80 (11) [29.99]	99.99 (15) [47.25]	-	-	99.99 (16) [292.78]	99.99 (11) [163.43]
2017	99.99 (16) [347.77]	99.99 (17) [85.96]	99.99 (19) [93.13]	99.99 (16) [100.18]	99.99 (11) [46.77]	99.99 (15) [655.18]	99.99 (16) [259.97]	99.90 (11) [30.36]	-	-

Table 4. Random parameters model results for single-vehicle single-pedestrian injury severity for Kansas for 2013-day data to (parameters defined for: [NVI] No visible injury; [MI] Moderate Injury; [SI] Severe Injury).

			Marginal effects		
Variable Description	**Estimated Parameter**	**t-statistic**	**No visible Injury**	**Moderate Injury**	**Severe Injury**
Constant [NVI]	0.98	2.05			
Constant [SI]	-2.21	-3.38			
Random parameter (normally distributed)					
Male driver indicator (1 if driver is male, 0 otherwise) [MI]	0.22	0.47	-0.0127	0.0193	-0.0066
Standard deviation of male driver indicator	*2.72*	*1.54*			
Vehicle characteristics					
Pickup vehicle indicator (1 if vehicle type was Pickup, 0 otherwise) [NVI]	-1.06	-2.26	-0.0336	0.0198	0.0138
SUV vehicle indicator (1 if vehicle type was SUV, 0 otherwise) [SI]	1.00	2.09	-0.0153	-0.0146	0.0299
Roadway characteristics					
Center or edge line indicator (1 if traffic control type was center or edge line, 0 otherwise) [NVI]	-0.70	-1.77	-0.0305	0.0176	0.0130
Crash characteristics					
Damaged vehicle indicator (1 if vehicle damaged was disabling, 0 otherwise) [SI]	2.26	2.42	-0.0062	-0.0067	0.0129
Vehicle side impact indicator (1 if vehicle principal impact is on the left side, 0 otherwise) [NVI]	1.79	2.59	0.0240	-0.0164	-0.0076
Driver characteristics					
Inattention on driving indicator (1 if driver was not paying attention, 0 otherwise) [NVI]	1.69	3.57	0.0530	-0.0316	-0.0214
Seatbelt indicator (1 if driver used seatbelt, 0 otherwise) [SI]	1.88	3.06	-0.1028	-0.0955	0.1984
Pedestrian characteristics					
No safety equipment indicator (1 if pedestrian did not use any pedestrian equipment, 0 otherwise) [NVI]	-0.85	-1.99	-0.1075	0.0670	0.0405
Younger pedestrian indicator (1 if pedestrian is less than 25 years old, 0 otherwise) [MI]	0.85	2.25	-0.0419	0.0707	-0.0288
Pedestrian signal disobedience indicator (1 if pedestrian disobeyed pedestrian signal, 0 otherwise) [SI]	2.55	2.90	-0.0093	-0.0126	0.0219
Model statistics					
Number of observations			227		
Log-likelihood at zero			-249.385		
Log-likelihood at convergence			-203.514		

Table 5. Random parameters model with heterogeneity in the means results for single-vehicle single-pedestrian injury severity for Kansas for 2013-night data to (parameters defined for: [NVI] No visible injury; [MI] Moderate Injury; [SI] Severe Injury).

			Marginal effects		
Variable Description	**Estimated Parameter**	**t-statistic**	**No visible Injury**	**Moderate Injury**	**Severe Injury**
Constant [NVI]	1.77	2.68			
Constant [SI]	2.37	2.70			
Random parameter (normally distributed)					
Roadway non-intersection indicator (1 if accident location was on roadway but non-intersection, 0 otherwise) [MI]	1.55	1.80	-0.0104	0.0226	-0.0122
Standard deviation of roadway non-intersection indicator	*2.49*	*1.24*			
Heterogeneity in the mean of random parameter					
Roadway non-intersection indicator (1 if accident location was on roadway but non-intersection, 0 otherwise); multiple lanes indicator (1 if the number of lanes is 3 or more, 0 otherwise) [MI]	-2.63	-1.91			
Roadway non-intersection indicator (1 if accident location was on roadway but non-intersection, 0 otherwise); younger driver indicator (1 if driver is less than 25 years old, 0 otherwise) [MI]	-2.70	-1.64			
Vehicle characteristics					
Pickup vehicle indicator (1 if vehicle type was Pickup, 0 otherwise) [NVI]	1.68	2.56	0.0477	-0.0135	-0.0342
Passenger vehicle indicator (1 if vehicle type was passenger vehicle, 0 otherwise) [SI]	-1.12	-2.10	0.0584	0.0415	-0.0999
Roadway characteristics					
Multiple lanes indicator (1 if the number of lanes is 3 or more, 0 otherwise) [NVI]	-1.38	-2.86	-0.0883	0.0408	0.0476
Center or edge line indicator (1 if traffic control type was center or edge line, 0 otherwise) [MI]	1.58	2.38	-0.0355	0.0953	-0.0598
Traffic signal indicator (1 if Traffic control type was signal, 0 otherwise) [SI]	-2.39	-2.62	0.0179	0.0115	-0.0295
No pedestrian signal indicator (1 if there was no pedestrian signal, 0 otherwise) [SI]	-1.36	-2.67	0.0691	0.0439	-0.113
Driver characteristics					
Male driver indicator (1 if driver is male, 0 otherwise) [SI]	0.99	2.08	-0.0405	-0.0347	0.0752
Pedestrian characteristics					
Younger pedestrian indicator (1 if pedestrian is less than 25 years old, 0 otherwise) [MI]	1.58	2.46	-0.0663	0.1035	-0.0372
Older pedestrian indicator (1 if pedestrian is 55 years old or older, 0 otherwise) [SI]	1.28	2.09	-0.0232	-0.0131	0.0363
Model statistics					
Number of observations			131		
Log-likelihood at zero			-143.918		
Log-likelihood at convergence			-114.743		

Table 6. Random parameters model with heterogeneity in the means results for single-vehicle single-pedestrian injury severity for Kansas for 2014-day data to (parameters defined for: [NVI] No visible injury; [MI] Moderate Injury; [SI] Severe Injury).

			Marginal effects		
Variable Description	**Estimated Parameter**	**t-statistic**	**No visible Injury**	**Moderate Injury**	**Severe Injury**
Constant [NVI]	2.31	2.60			
Constant [SI]	-0.83	-1.00			
Random parameter (normally distributed)					
Passenger vehicle indicator (1 if vehicle type was passenger vehicle, 0 otherwise) [MI]	-0.08	-0.10	0.0303	-0.038	0.0077
Standard deviation of passenger vehicle indicator	*3.59*	*1.77*			
Heterogeneity in the mean of random parameter					
Passenger vehicle indicator (1 if vehicle type was passenger vehicle, 0 otherwise); center or edge line indicator (1 if traffic control type was center or edge line, 0 otherwise) [MI]	-2.01	-1.52			
Passenger vehicle indicator (1 if vehicle type was passenger vehicle, 0 otherwise); female pedestrian indicator (1 if pedestrian is female, 0 otherwise) [MI]	-1.66	-1.36			
Vehicle characteristics					
Pickup vehicle indicator (1 if vehicle type was Pickup, 0 otherwise) [NVI]	0.89	1.68	0.0277	-0.0194	-0.0084
SUV vehicle indicator (1 if vehicle type was SUV, 0 otherwise) [SI]	2.44	3.58	-0.0175	-0.0268	0.0442
Pickup vehicle indicator (1 if vehicle type was Pickup, 0 otherwise) [SI]	2.03	2.67	-0.0191	-0.0144	0.0334
Roadway characteristics					
Dry pavement indicator (1 if road surface condition is dry, 0 otherwise) [NVI]	-1.18	-1.75	-0.1591	0.1115	0.0475
Low road speed limit indicator (1 if road speed limit was 30 m/h or less, 0 otherwise) [SI]	-1.31	-2.45	0.0316	0.0203	-0.0519
Road surface characteristic indicator (1 if the road was straight and level, 0 otherwise) [SI]	-1.27	-1.99	0.0419	0.0267	-0.0686
Crash characteristics					
Vehicle side impact indicator (1 if vehicle principal impact is on the left side, 0 otherwise) [NVI]	-3.48	-2.74	-0.0165	0.0092	0.0072
Vehicle front impact indicator (1 if vehicle principal impact was on the front, 0 otherwise) [SI]	1.12	3.65	-0.0656	-0.0384	0.1040
Pedestrian location indicator (1 if pedestrian was not in roadway, 0 otherwise) [SI]	-2.73	-2.34	0.0048	0.0047	-0.0095
Driver characteristics					
Driver action indicator (1 if driver failed to yield the right of way, 0 otherwise) [SI]	-1.75	-2.43	0.0112	0.0094	-0.0206
Younger driver indicator (1 if driver is less than 25 years old, 0 otherwise) [SI]	-1.43	-2.37	0.0155	0.0086	-0.0241
Pedestrian characteristics					
No safety equipment indicator (1 if pedestrian did not use any pedestrian equipment, 0 otherwise) [NVI]	-1.52	-2.57	-0.1923	0.1332	0.0590
Younger pedestrian indicator (1 if pedestrian is less than 25 years old, 0 otherwise) [MI]	0.99	2.23	-0.0433	0.0542	-0.0109
Model statistics					
Number of observations			229		
Log-likelihood at zero			-251.582		
Log-likelihood at convergence			-182.214		

Table 7. Random parameters model results for single-vehicle single-pedestrian injury severity for Kansas for 2014-night data to (parameters defined for: [NVI] No visible injury; [MI] Moderate Injury; [SI] Severe Injury).

			Marginal effects		
Variable Description	**Estimated Parameter**	**t-statistic**	**No visible Injury**	**Moderate Injury**	**Severe Injury**
Constant [NVI]	0.86	1.53			
Constant [SI]	-3.72	-3.64			
Random parameter (normally distributed)					
Warmer-weather months indicator (1 if accident occurs from March until October, 0 otherwise) [NVI]	0.91	1.43	0.0551	-0.0312	-0.0239
Standard deviation of warmer-weather months indicator	*3.55*	*1.89*			
Vehicle characteristics					
Van vehicle indicator (1 if vehicle type was Van, 0 otherwise) [NVI]	3.27	2.68	0.0200	-0.0129	-0.0071
Roadway characteristics					
Wet pavement indicator (1 if road surface condition is wet, 0 otherwise) [MI]	-1.48	-1.73	-0.022	0.0138	0.0082
Flexible pavement indicator (1 if road surface type is asphalt, 0 otherwise) [SI]	1.04	1.85	-0.0304	-0.0515	0.0819
Crash characteristics					
Roadway non-intersection indicator (1 if accident location was on roadway but non-intersection, 0 otherwise) [NVI]	-1.37	-2.17	-0.0634	0.0356	0.0278
Vehicle maneuver indicator (1 if vehicle maneuver was straight and following the road, 0 otherwise) [SI]	1.71	2.98	-0.0522	-0.0868	0.139
Driver characteristics					
Not-injured driver indicator (1 if driver was not injured, 0 otherwise) [SI]	1.63	2.17	-0.0561	-0.0947	0.1507
Pedestrian characteristics					
Male pedestrian indicator (1 if pedestrian is Male, 0 otherwise) [NVI]	-1.52	-2.52	-0.1093	0.0691	0.0402
Running pedestrian indicator (1 if pedestrian action was running or playing, 0 otherwise) [MI]	-1.74	-3.25	0.0326	-0.0692	0.0366
Pedestrian signal disobedience indicator (1 if pedestrian disobeyed pedestrian signal, 0 otherwise) [MI]	3.89	2.16	-0.0111	0.0142	-0.0031
Younger pedestrian indicator (1 if pedestrian is less than 25 years old, 0 otherwise) [SI]	-1.63	-2.72	0.0112	0.0250	-0.0362
Model statistics					
Number of observations			161		
Log-likelihood at zero			-176.876		
Log-likelihood at convergence			-133.233		

Table 8. Multinomial logit model results (random parameters were not statistically significant) for single-vehicle single-pedestrian injury severity for Kansas for 2015-day data to (parameters defined for: [NVI] No visible injury; [MI] Moderate Injury; [SI] Severe Injury).

			Marginal effects		
Variable Description	**Estimated Parameter**	**t-statistic**	**No visible Injury**	**Moderate Injury**	**Severe Injury**
Constant [NVI]	0.84	3.00			
Constant [SI]	-0.85	-2.14			
Vehicle characteristics					
SUV vehicle indicator (1 if vehicle type was SUV, 0 otherwise) [NVI]	-1.27	-2.90	-0.2772	0.1873	0.0899
Roadway characteristics					
No pedestrian signal indicator (1 if there was no pedestrian signal, 0 otherwise) [NVI]	-0.79	-2.35	-0.1724	0.1164	0.0559
Wet pavement indicator (1 if road surface condition is wet, 0 otherwise) [NVI]	1.43	1.85	0.3127	-0.2112	-0.1015
Traffic signal indicator (1 if Traffic control type was signal, 0 otherwise) [MI]	1.21	3.04	-0.1794	0.2652	-0.0858
Crash characteristics					
Vehicle front impact indicator (1 if vehicle principal impact was on the front, 0 otherwise) [SI]	0.91	2.27	-0.0647	-0.0644	0.1292
Roadway non-intersection indicator (1 if accident location was on roadway but non-intersection, 0 otherwise) [SI]	0.78	1.91	-0.0552	-0.0550	0.1102
Pedestrian characteristics					
Younger pedestrian indicator (1 if pedestrian is less than 25 years old, 0 otherwise) [SI]	-0.83	-2.05	0.0587	0.0585	-0.1172
Model statistics					
Number of observations			188		
Log-likelihood at zero			-197.160		
Log-likelihood at convergence			-180.637		

Table 9. Random parameters model results for single-vehicle single-pedestrian injury severity for Kansas for 2015-night data to (parameters defined for: [NVI] No visible injury; [MI] Moderate Injury; [SI] Severe Injury).

Variable Description	Estimated Parameter	t-statistic	Marginal effects: No visible Injury	Marginal effects: Moderate Injury	Marginal effects: Severe Injury
Constant [NVI]	-0.96	-2.26			
Constant [SI]	-2.00	-2.99			
Random parameter (normally distributed)					
Younger driver indicator (1 if driver is less than 25 years old, 0 otherwise) [SI]	-4.08	-1.25	0.0019	0.0021	-0.0041
Standard deviation of younger driver indicator	*3.92*	*1.27*			
Vehicle characteristics					
SUV vehicle indicator (1 if vehicle type was SUV, 0 otherwise) [MI]	1.28	2.57	-0.0318	0.0442	-0.0124
Roadway characteristics					
Multiple lanes indicator (1 if the number of lanes is 3 or more, 0 otherwise) [NVI]	-1.38	-2.97	-0.0835	0.0559	0.0276
Traffic signal indicator (1 if Traffic control type was signal, 0 otherwise) [MI]	-2.36	-3.32	0.0428	-0.0561	0.0132
No pedestrian signal indicator (1 if there was no pedestrian signal, 0 otherwise) [MI]	-1.07	-2.51	0.0728	-0.1155	0.0427
Stop sign indicator (1 if Traffic control type was stop sign, 0 otherwise) [SI]	3.31	2.54	-0.0201	-0.0083	0.0284
Center or edge line indicator (1 if traffic control type was center or edge line, 0 otherwise) [SI]	1.42	2.46	-0.0317	-0.063	0.0947
Low road speed limit indicator (1 if road speed limit was 30 m/h or less, 0 otherwise) [SI]	-1.07	-1.95	0.0237	0.0240	-0.0477
Crash characteristics					
Roadway non-intersection indicator (1 if accident location was on roadway but non-intersection, 0 otherwise) [SI]	1.25	2.39	-0.0288	-0.0362	0.0650
Driver characteristics					
Alcohol-influenced driver indicator (1 if the driver was under the influence of alcohol, 0 otherwise) [MI]	-1.92	-1.64	0.0077	-0.0093	0.0017
Pedestrian characteristics					
Walking pedestrian indicator (1 if pedestrian action was walking, 0 otherwise) [NVI]	1.51	2.46	0.0374	-0.0241	-0.0134
Younger pedestrian indicator (1 if pedestrian is less than 25 years old, 0 otherwise) [NVI]	0.83	2.05	0.0632	-0.0439	-0.0192
Model statistics					
Number of observations			154		
Log-likelihood at zero			-169.186		
Log-likelihood at convergence			-132.945		

Table 10. Random parameters model with heterogeneity in the means results for single-vehicle single-pedestrian injury severity for Kansas for 2016-day data to (parameters defined for: [NVI] No visible injury; [MI] Moderate Injury; [SI] Severe Injury).

Variable Description	Estimated Parameter	t-statistic	Marginal effects: No visible Injury	Marginal effects: Moderate Injury	Marginal effects: Severe Injury
Constant [NVI]	1.07	1.06			
Constant [SI]	0.85	0.78			
Random parameter (normally distributed)					
Dry pavement indicator (1 if road surface condition is dry, 0 otherwise) [MI]	-2.39	-1.64	-0.0237	0.0227	0.0010
Standard deviation of dry pavement indicator	*5.42*	*2.28*			
Heterogeneity in the mean of random parameter					
Dry pavement indicator (1 if road surface condition is dry, 0 otherwise); weekend indicator (1 if accident occurred on the weekend days, 0 otherwise) [MI]	3.34	1.98			
Dry pavement indicator (1 if road surface condition is dry, 0 otherwise): younger pedestrian indicator (1 if pedestrian is less than 25 years old, 0 otherwise) [MI]	1.64	1.42			
Temporal characteristics					
Warmer-weather months indicator (1 if accident occurs from March until October, 0 otherwise) [MI]	1.49	1.69	-0.0535	0.0731	-0.0196
Vehicle characteristics					
SUV vehicle indicator (1 if vehicle type was SUV, 0 otherwise) [NVI]	-1.45	-2.70	-0.0376	0.0175	0.0201
Pickup vehicle indicator (1 if vehicle type was Pickup, 0 otherwise) [SI]	1.89	3.64	-0.0350	-0.0112	0.0462
Roadway characteristics					
Low road speed limit indicator (1 if road speed limit was 30 m/h or less, 0 otherwise) [NVI]	1.32	2.90	0.0971	-0.0469	-0.0502
Road surface characteristic indicator (1 if the road was straight and level, 0 otherwise) [SI]	-1.11	-1.98	0.0560	0.0159	-0.0718
Crash characteristics					
Vehicle maneuver indicator (1 if vehicle maneuver was a right turn, 0 otherwise) [NVI]	3.68	2.64	0.0285	-0.0231	-0.0054
Vehicle front impact indicator (1 if vehicle principal impact was on the front, 0 otherwise) [MI]	2.38	2.43	-0.0445	0.0607	-0.0162
Pedestrian location indicator (1 if pedestrian was not in roadway, 0 otherwise) [SI]	-1.88	-1.67	0.0054	0.0015	-0.0069
Driver characteristics					
Driver action indicator (1 if the driver disregarded traffic signs, 0 otherwise) [NVI]	6.88	1.66	0.0066	-0.0065	-0.0001
Pedestrian characteristics					
Older pedestrian indicator (1 if pedestrian is 55 years old or older, 0 otherwise) [SI]	1.52	3.03	-0.0333	-0.0099	0.0432
Model statistics					
Number of observations			264		
Log-likelihood at zero			-290.034		
Log-likelihood at convergence			-219.831		

Table 11. Random parameters model with heterogeneity in the means results for single-vehicle single-pedestrian injury severity for Kansas for 2016-night data to (parameters defined for: [NVI] No visible injury; [MI] Moderate Injury; [SI] Severe Injury).

Variable Description	Estimated Parameter	t-statistic	Marginal effects: No visible Injury	Marginal effects: Moderate Injury	Marginal effects: Severe Injury
Constant [NVI]	0.36	0.63			
Constant [SI]	1.68	2.12			
Random parameter (normally distributed)					
Dry pavement indicator (1 if road surface condition is dry, 0 otherwise) [MI]	-5.04	-1.46	-0.0008	0.0206	-0.0198
Standard deviation of dry pavement indicator	*4.77*	*1.68*			
Heterogeneity in the mean of random parameter					
Dry pavement indicator (1 if road surface condition is dry, 0 otherwise); male pedestrian indicator (1 if pedestrian is male, 0 otherwise) [MI]	2.15	1.31			
Dry pavement indicator (1 if road surface condition is dry, 0 otherwise): no safety equipment indicator (1 if pedestrian did not use any pedestrian equipment, 0 otherwise) [MI]	3.46	1.32			
Temporal characteristics					
Weekend indicator (1 if accident occurred on the weekend days, 0 otherwise) [SI]	1.65	2.48	-0.0383	-0.0322	0.0706
Vehicle characteristics					
Passenger vehicle indicator (1 if vehicle type was passenger vehicle, 0 otherwise) [SI]	-2.08	-3.17	0.0692	0.041	-0.1101
Roadway characteristics					
Low road speed limit indicator (1 if road speed limit was 30 m/h or less, 0 otherwise) [SI]	-2.44	-3.31	0.0566	0.0436	-0.1002
Intersection type indicator (1 if intersection type is four way, 0 otherwise) [SI]	-2.26	-2.87	0.0249	0.0132	-0.0381
Crash characteristics					
Vehicle front impact indicator (1 if vehicle principal impact was on the front, 0 otherwise) [NVI]	-2.91	-3.61	-0.0651	0.0201	0.0449
Pedestrian location indicator (1 if pedestrian was not in roadway, 0 otherwise) [SI]	-2.57	-2.61	0.0244	0.0074	-0.0318
Pedestrian characteristics					
Older pedestrian indicator (1 if pedestrian is 55 years old or older, 0 otherwise) [SI]	1.26	2.00	-0.0219	-0.0205	0.0424
Model statistics					
Number of observations			151		
Log-likelihood at zero			-165.890		
Log-likelihood at convergence			-128.699		

Table 12. Multinomial logit model results (random parameters were not statistically significant) results for single-vehicle single-pedestrian injury severity for Kansas for 2017-day data to (parameters defined for: [NVI] No visible injury; [MI] Moderate Injury; [SI] Severe Injury).

			Marginal effects		
Variable Description	**Estimated Parameter**	**t-statistic**	**No visible Injury**	**Moderate Injury**	**Severe Injury**
Constant [NVI]	2.16	2.23			
Constant [SI]	-1.92	-2.47			
Temporal characteristics					
Weekend indicator (1 if accident occurred on the weekend days, 0 otherwise) [MI]	0.58	1.81	-0.0861	0.1242	-0.0381
Warmer-weather months indicator (1 if accident occurs from March until October, 0 otherwise) [SI]	1.37	2.35	-0.0741	-0.0907	0.1647
Vehicle characteristics					
Pickup vehicle indicator (1 if vehicle type was Pickup, 0 otherwise) [MI]	0.99	2.37	-0.1484	0.2139	-0.0656
Roadway characteristics					
Multiple lanes indicator (1 if the number of lanes is 3 or more, 0 otherwise) [NVI]	-1.60	-3.91	-0.3249	0.2387	0.0862
No pedestrian signal indicator (1 if there was no pedestrian signal, 0 otherwise) [NVI]	-0.96	-2.49	-0.1957	0.1438	0.0519
Weather condition indicator (1 if there are no adverse conditions, 0 otherwise) [NVI]	-1.47	-1.74	-0.2984	0.2193	0.0792
Traffic signal indicator (1 if Traffic control type was signal, 0 otherwise) [NVI]	1.07	2.42	0.2181	-0.1603	-0.0579
Low road speed limit indicator (1 if road speed limit was 30 m/h or less, 0 otherwise) [SI]	-1.14	-2.78	0.0618	0.0757	-0.1374
Crash characteristics					
Vehicle side impact indicator (1 if vehicle principal impact is on the right or left side, 0 otherwise) [NVI]	0.79	2.00	0.1610	-0.1183	-0.0427
Driver characteristics					
Male driver indicator (1 if driver is male, 0 otherwise) [NVI]	0.75	2.26	0.1528	-0.1123	-0.0405
Seatbelt indicator (1 if driver used seatbelt, 0 otherwise) [SI]	1.08	1.83	-0.0580	-0.0711	0.1291
Model statistics					
Number of observations			203		
Log-likelihood at zero			-207.840		
Log-likelihood at convergence			-181.310		

Table 13. Random parameters model with heterogeneity in the means results for single-vehicle single-pedestrian injury severity for Kansas for 2017-night data to (parameters defined for: [NVI] No visible injury; [MI] Moderate Injury; [SI] Severe Injury).

			Marginal effects		
Variable Description	**Estimated Parameter**	**t-statistic**	**No visible Injury**	**Moderate Injury**	**Severe Injury**
Constant [NVI]	0.03	0.03			
Constant [SI]	0.74	1.24			
Random parameter (normally distributed)					
Warmer-weather months indicator (1 if accident occurs from March until October, 0 otherwise) [MI]	0.21	0.26	-0.0191	0.0486	-0.0294
Standard deviation of warmer-weather indicator	*3.36*	*1.84*			
Heterogeneity in the mean of random parameter					
Warmer-weather months indicator (1 if accident occurs from March until October, 0 otherwise); running pedestrian indicator (1 if pedestrian action was running or playing, 0 otherwise) [MI]	2.05	1.38			
Warmer-weather months indicator (1 if accident occurs from March until October, 0 otherwise); walking pedestrian indicator (1 if pedestrian action was walking, 0 otherwise) [MI]	-3.31	-1.21			
Vehicle characteristics					
Pickup vehicle indicator (1 if vehicle type was Pickup, 0 otherwise) [NVI]	-1.33	-2.08	-0.029	0.0121	0.0169
Roadway characteristics					
Low road speed limit indicator (1 if road speed limit was 30 m/h or less, 0 otherwise) [MI]	2.31	3.33	-0.0938	0.1612	-0.0674
Intersection type indicator (1 if intersection type was T-intersection, 0 otherwise) [MI]	-2.03	-1.78	0.0088	-0.0158	0.0070
Crash characteristics					
Damaged vehicle indicator (1 if vehicle damaged was disabling, 0 otherwise) [SI]	3.51	3.10	-0.0129	-0.0156	0.0286
Vehicle front impact indicator (1 if vehicle principal impact was on the front, 0 otherwise) [SI]	1.25	2.56	-0.0557	-0.0371	0.0928
Driver characteristics					
Younger driver indicator (1 if driver is less than 25 years old, 0 otherwise) [SI]	-1.31	-2.39	0.0316	0.0174	-0.0490
Pedestrian characteristics					
No safety equipment indicator (1 if pedestrian did not use any pedestrian equipment, 0 otherwise) [NVI]	1.65	1.79	0.2368	-0.1033	-0.1336
Model statistics					
Number of observations			141		
Log-likelihood at zero			-154.904		
Log-likelihood at convergence			-124.151		

Table 14. Comparison of marginal effects of single-pedestrian single-vehicle crashes over the years for daytime, and for nighttime in parentheses.

Variables	No visible Injury					Moderate Injury					Severe Injury				
	2013	2014	2015	2016	2017	2013	2014	2015	2016	2017	2013	2014	2015	2016	2017
Temporal characteristics															
Warmer-weather months indicator (1 if accident occurs from March until October, 0 otherwise)	-	-	-	-0.0535	-0.0741	-	-	-	0.0731	-0.0907	-	-	-	-0.0196	0.1647
	-	(0.0551)	-	-	(-0.0191)	-	(-0.0312)	-	-	(0.0486)	-	(-0.0239)	-	-	(-0.0294)
Weekend indicator (1 if accident occurred on the weekend days, 0 otherwise)	-	-	-		-0.0861	-	-	-	-	0.1242	-	-	-	-	-0.0381
	-	-	-	(-0.0383)	-	-	-	-	(-0.0322)	-	-	-	-	(0.0706)	-
Vehicle characteristics															
Pickup vehicle indicator (1 if vehicle type was Pickup, 0 otherwise)	-0.0336	0.0277	-	-0.0350	-0.1484	0.0198	-0.0194	-	-0.0112	0.2139	0.0138	-0.0084	-	0.0462	-0.0656
	(0.0477)	-0.0191	-	-	(-0.0290)	(-0.0135)	-0.0144	-	-	(0.0121)	(-0.0342)	0.0334	-	-	(0.0169)
SUV vehicle indicator (1 if vehicle type was SUV, 0 otherwise)	-0.0153	-0.0175	-0.2772	-0.0376	-	-0.0146	-0.0268	0.1873	0.0175	-	0.0299	0.0442	0.0900	0.0201	-
	-	-	(-0.0318)	-	-	-	-	(0.0442)	-	-	-	-	(-0.0124)	-	-
Passenger vehicle indicator (1 if vehicle type was passenger vehicle, 0 otherwise)	-	0.0303	-	-	-	-	-0.0380	-	-	-	-	0.0077	-	-	-
	(-0.0584)	-	-	(0.0692)	-	(-0.0415)	-	-	(0.0410)	-	(-0.0999)	-	-	(-0.1101)	-
Van vehicle indicator (1 if vehicle type was Van, 0 otherwise)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	-	(0.0200)	-	-	-	-	(-0.0129)	-	-	-	-	(-0.0071)	-	-	-
Roadway characteristics															
Center or edge line indicator (1 if traffic control type was center or edge line, 0 otherwise)	-0.0305	-	-	-	-	0.0176	-	-	-	-	0.013	-	-	-	-
	(-0.0355)	-	(-0.0317)	-	-	(0.0953)	-	(-0.0630)	-	-	(-0.0598)	-	(0.0947)	-	-
Stop sign indicator (1 if Traffic control type was stop sign, 0 otherwise)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	-	-	(-0.0201)	-	-	-	-	(-0.0083)	-	-	-	-	(0.0284)	-	-
Traffic signal indicator (1 if Traffic control type was signal, 0 otherwise)	-	-	-0.1794	-	0.2181	-	-	0.2652	-	-0.1603	-	-	-0.0858	-	-0.0579
	(0.0179)	-	(0.0428)	-	-	(0.0115)	-	(-0.0561)	-	-	(-0.0295)	-	(0.0132)	-	-
No pedestrian signal indicator (1 if there was no pedestrian signal, 0 otherwise)	-	-	-0.1724	-	-0.1957	-	-	0.1164	-	0.1438	-	-	0.0559	-	0.0519
	(0.0691)	-	(0.0728)	-	-	(0.0439)	-	(-0.1155)	-	-	(-0.1130)	-	(0.0427)	-	-
Low road speed limit indicator (1 if road speed limit was 30 m/h or less , 0 otherwise)	-	0.0316	-	0.0971	0.0618	-	0.0203	-	-0.0469	0.0757	-	-0.0519	-	-0.0502	-0.1374
	-	-	(0.0237)	(0.0566)	(-0.0938)	-	-	(0.0240)	(0.0436)	(0.1612)	-	-	(-0.0477)	(-0.1002)	(-0.0674)
Intersection type indicator (1 if intersection type is four way, 0 otherwise)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	-	-	-	(0.0249)	-	-	-	-	(0.0132)	-	-	-	-	(-0.0381)	-
Intersection type indicator (1 if intersection type was T-intersection, 0 otherwise)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	-	-	-	-	(0.0088)	-	-	-	-	(-0.0158)	-	-	-	-	(0.0070)
Multiple lanes indicator (1 if the number of lanes is 3 or more, 0 otherwise)	-	-	-	-	-0.3249	-	-	-	-	0.2387	-	-	-	-	0.0862
	(-0.0883)	-	(-0.0835)	-	-	(0.0408)	-	(0.0559)	-	-	(0.0476)	-	(0.0276)	-	-
Road surface characteristic indicator (1 if the road was straight and level, 0 otherwise)	-	0.0419	-	0.056	-	-	0.0267	-	0.0159	-	-	-0.0686	-	-0.0718	-
	-		-		-	-		-		-	-		-		-
Dry pavement indicator (1 if road surface condition is dry, 0 otherwise)	-	-0.1591	-	-0.0237	-	-	0.1115	-	0.0227	-	-	0.0475	-	0.0010	-
	-	-	-	(-0.0008)	-	-	-	-	(0.0206)	-	-	-	-	(-0.0198)	-
Wet pavement indicator (1 if road surface condition is wet, 0 otherwise)	-	-	0.3127	-	-	-	-	-0.2112	-	-	-	-	-0.1015	-	-
	-	(-0.0220)	-	-	-	-	(0.0138)	-	-	-	-	(0.0082)	-	-	-
Flexible pavement indicator (1 if road surface type is asphalt, 0 otherwise)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	-	(-0.0304)	-	-	-	-	(-0.0515)	-	-	-	-	(0.0819)	-	-	-
Weather condition indicator (1 if there are no adverse conditions, 0 otherwise)	-	-	-	-	-0.2984	-	-	-	-	0.2193	-	-	-	-	0.0792
	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Crash characteristics															
Damaged vehicle indicator (1 if vehicle damaged was disabling, 0 otherwise)	-0.0062	-	-	-	-	-0.0067	-	-	-	-	0.0129	-	-	-	-
	-	-	-	-	(-0.0129)	-	-	-	-	(-0.0156)	-	-	-	-	(0.0286)

Vehicle front impact indicator (1 if vehicle principal impact was on the front, 0 otherwise)	-	-0.0656	-0.0647	-0.0445	-	-	-0.0384	-0.0644	0.0607	-	-	0.1040	0.1292	-0.0162	-
	-	-	-	(-0.0651)	(-0.0557)	-	-	-	(0.0201)	(-0.0370)	-	-	-	(0.0449)	(0.0928)
Roadway non-intersection indicator (1 if accident location was on roadway but non-intersection, 0 otherwise)	-	-	-0.0552	-	-	-	-	-0.0550	-	-	-	-	0.1102	-	-
	(-0.0104)	(-0.0634)	(-0.0288)	-	-	(0.0226)	(0.0356)	(-0.0362)	-	-	(-0.0122)	(0.0278)	(0.0650)	-	-
Pedestrian location indicator (1 if pedestrian was not in roadway, 0 otherwise)	-	0.0048	-	0.0054	-	-	0.0047	-	0.0015	-	-	-0.0095	-	-0.0069	-
	-	-	-	(0.0244)	-	-	-	-	(0.0074)	-	-	-	-	(-0.0318)	-
Vehicle side impact indicator (1 if vehicle principal impact is on the left side, 0 otherwise)	0.0240	-0.0165	-	-	-	-0.0164	0.0092	-	-	-	-0.0076	0.0072	-	-	-
	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Vehicle side impact indicator (1 if vehicle principal impact is on the right or left side, 0 otherwise)	-	-	-	-	0.1610	-	-	-	-	-0.1183	-	-	-	-	-0.0427
	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Vehicle maneuver indicator (1 if vehicle maneuver was straight and following the road, 0 otherwise)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	-	(-0.0522)	-	-	-	-	(-0.0868)	-	-	-	-	(0.1390)	-	-	-
Vehicle maneuver indicator (1 if vehicle maneuver was a right turn, 0 otherwise)	-	-	-	0.0285	-	-	-	-	-0.0231	-	-	-	-	-0.0054	-
	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Driver characteristics															
Inattention on driving indicator (1 if driver was not paying attention, 0 otherwise)	0.0530	-	-	-	-	-0.0316	-	-	-	-	-0.0214	-	-	-	-
	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Seatbelt indicator (1 if driver used seatbelt, 0 otherwise)	-0.1028	-	-	-	-0.0580	-0.0955	-	-	-	-0.0711	0.1984	-	-	-	0.1291
	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Driver action indicator (1 if driver failed to yield the right of way, 0 otherwise)	-	0.0112	-	-	-	-	0.0094	-	-	-	-	-0.0206	-	-	-
	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Younger driver indicator (1 if driver is less than 25 years old, 0 otherwise)	-	0.0155	-	-	-	-	0.0086	-	-	-	-	-0.0241	-	-	-
	-	-	(0.0019)	-	(0.0316)	-	-	(0.0021)	-	(0.0174)	-	-	(-0.0041)	-	(-0.0490)
Driver action indicator (1 if the driver disregarded traffic signs, 0 otherwise)	-	-	-	0.0066	-	-	-	-	-0.0065	-	-	-	-	-0.0001	-
	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Male driver indicator (1 if driver is male, 0 otherwise)	-0.0127	-	-	-	0.1528	0.0193	-	-	-	-0.1123	-0.0066	-	-	-	-0.0405
	(-0.0405)	-	-	-	-	(-0.0347)	-	-	-	-	(0.0752)	-	-	-	-
Not-injured driver indicator (1 if driver was not injured, 0 otherwise)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	-	(-0.0561)	-	-	-	-	-(-0.0947)	-	-	-	-	(0.1507)	-	-	-
Alcohol-influenced driver indicator (1 if the driver was under the influence of alcohol, 0 otherwise)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	-	-	(0.0077)	-	-	-	-	(-0.0093)	-	-	-	-	(0.0017)	-	-
Pedestrian characteristics															
Younger pedestrian indicator (1 if pedestrian is less than 25 years old, 0 otherwise)	-0.0419	-0.0433	0.0587	-	-	0.0707	0.0542	0.0585	-	-	-0.0288	-0.0109	-0.1172	-	-
	(-0.0663)	(0.0112)	(0.0632)	-	-	(0.1035)	(0.0250)	(-0.0439)	-	-	(-0.0372)	(-0.0362)	(-0.0192)	-	-
Older pedestrian indicator (1 if pedestrian is 55 years old or older, 0 otherwise)	-	-	-	-0.0333	-	-	-	-	-0.0099	-	-	-	-	0.0432	-
	(-0.0232)	-	-	(-0.0219)	-	(-0.0131)	-	-	(-0.0205)	-	(0.0363)	-	-	(0.0424)	-
No safety equipment indicator (1 if pedestrian did not use any pedestrian equipment, 0 otherwise)	-0.1075	-0.1923	-	-	-	0.067	0.1332	-	-	-	0.0405	0.0590	-	-	-
	-	-	-	-	(0.2368)	-	-	-	-	(-0.1033)	-	-	-	-	(-0.1336)
Pedestrian signal disobedience indicator (1 if pedestrian disobeyed pedestrian signal, 0 otherwise)	-0.0093	-	-	-	-	-0.0126	-	-	-	-	0.0219	-	-	-	-
	-	(-0.0111)	-	-	-	-	(0.0142)	-	-	-	-	(-0.0031)	-	-	-
Male pedestrian indicator (1 if pedestrian is male, 0 otherwise)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	-	(-0.1093)	-	-	-	-	(0.0691)	-	-	-	-	(0.0402)	-	-	-
Running pedestrian indicator (1 if pedestrian action was running or playing, 0 otherwise)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	-	(0.0326)	-	-	-	-	(-0.0692)	-	-	-	-	(0.0366)	-	-	-
Walking pedestrian indicator (1 if pedestrian action was walking, 0 otherwise)	-	-		-	-	-	-	-	-	-	-	-	-	-	-
	-	-	-(0.0374)	-	-	-	-	(-0.0241)	-	-	-	-	(-0.0134)	-	-

included more than one random parameter did not result in statistically significant correlation among random parameters. Final model estimations did not include any model with more than one random parameter.

Starting first with variables that produced statistically significant random parameters, Tables 4-13 show that out of all eight models had one statistically significant random parameter with the explanatory variable varying by year and time of day. Out of these random parameter variables, five produced statistically significant heterogeneity in the mean (again, none had a statistically significant heterogeneity in the variance). Estimates in Tables 4 and 5 show that the male driver indicator and the roadway non-intersection indicator (accidents occur on the roadway but not at an intersection) produced statistically significant random parameters in the 2013-daytime and 2013-nighttime models, respectively. Marginal effects in Table 14 show that the average effects of both of these indicators increased the likelihood of pedestrian moderate injury (and decreased the likelihoods of no visible injury and severe injury). The roadway non-intersection indicator, however, had statistically significant heterogeneity in the means, with the effect of accidents occurring on roadway but not at an intersection varying by multiple lanes (3 lanes or more) and younger drivers (less than 25 years old), with both of these deceasing the mean of the random parameter.

Estimations results for 2014-day accidents (Table 6) indicate that the passenger-vehicle indicator produced a random parameter (the effect of this variable on injury severities varied across accident observations) with the average marginal effects of this indicator decreasing the likelihood of moderate injury while increasing the likelihoods of no visible injury and severe injury. This variable also had statistically significant heterogeneity in the means, with the effect of accidents involved passenger vehicle varying by control type indicator (center or edge line) and female pedestrian both of these deceasing the mean of the random parameter.

Model results in Tables 7 and 13 show that the warmer-weather months indicator (March through October) produced statistically significant random parameters in both 2014 and 2017 nighttime models capturing an unobserved heterogeneity in accidents occur at this time of day. Marginal effects in Table 14, however, show that this variable had different average effects between the two years increasing the likelihood of no visible injury in 2014 and the likelihood of pedestrian moderate injury in 2017 (and correspondingly decreasing the other two injury-severity outcomes). Additionally, this variable had statistically significant heterogeneity in the means in the 2017 nighttime model, with the effect of accidents occurring during warmer-weather months varying by running and walking pedestrians. In this case, running increased the parameter mean resulting in a higher likelihood of moderate injury and walking decreased the parameter mean resulting in a lower likelihood of moderate injury.

Estimates in Table 9 show that the younger driver indicator (less than 25 years old) produced a statistically significant random parameter in the 2015 nighttime model, with the overall effect of this variable decreasing the likelihood of severe injury (as indicated by marginal effects). Estimation results in Tables 10 and 11 show that the dry pavement indicator resulted in a statistically significant random parameter variable in both of the 2016 models (daytime and nighttime). Marginal effects show that the average effect of this variable in both models was to increase the likelihood of moderate injury and decrease the likelihood of no visible injury. The presence of different random parameters in different years and times of day could be associated with period-specific unobserved effects (Behnood and Mannering, 2016).

Turning to variables that produced statistically significant parameters that were fixed across observations, Tables 4-13 show a variety of variables related to temporal, vehicle, roadway, crash, weather, and pedestrian-related characteristics produced statistically significant random parameters. The effects of these variables are discussed below and as previously mentioned, Table 14 shows a comparison of marginal effects for all variables in different years and times of day.

1.6.1 Time- and weather-related characteristics

Results in Table 12 show that the no-adverse-weather condition indicator was significant in one model (2017-daytime) with the average effect being a decrease in no visible injury, which interestingly contradicts the earlier finding of Forbes and Habib (2015). The weekend indicator variable produced statistically significant parameters in two models, 2016 nighttime and 2017 daytime, with the average effect of this variable being somewhat different in the two models; increasing the likelihood of pedestrian severe injury in the 2016-nighttime model and the likelihood of pedestrian moderate injury in the 2017-daytime model (see Table 14). Tables 10 and 12 show that the warmer-weather months indicator (March through October) was a statistically significant factor in the daytime in two consecutive years, 2016 and 2017, with the average effect being a decrease in no visible injury (with an increase in moderate injury in 2016 and an increase in severe injury in 2017), illustrating different marginal effects results relative to the random parameters nighttime models which found this variable to be statistically insignificant at night in 2016 and resulting in a random parameter at night in 2017.

1.6.2 Vehicle-related characteristics

As shown in Tables 4 through 13, different types of vehicles resulted in different impacts on pedestrian injury severity over the years depending on time of day. For example, the pickup-truck indicator vehicle was a statistically significant variable in all daytime models (except for 2015), showing an average effect of no visible injury being decreased with varied marginal effects magnitudes over the years. However, when this variable was defined for severe injury in 2014-daytime model, it showed an opposite effect of no visible injury being increased (see Table 14). For the nighttime models, the pickup truck indicator was statistically significant in two years (2013 and 2017) with an opposite average effect of severe injury being decreased in 2013 and increased in 2017. The Sports Utility Vehicle (SUV) indicator was a statistically significant variable in all daytime models (except for 2017) showing a relatively consistent direction but different magnitudes of the average marginal effects over the years. Marginal effects in Table 14 show that SUVs increased the probability of severe injury and decreased the probability of no visible injury in all daytime models except 2017. For nighttime, SUV vehicle type was a statistically significant variable in one year (2015) with a slightly different average effect being an increase in moderate injury and a decrease in no visible and severe injuries. Clearly, larger vehicle types (Pickups and SUVs) are generally found to result in higher injury severity levels, likely due to heavier mass and higher frontal designs. These findings confirm the results of previous work by Pour-Rouholamin and Zhou (2016) and Mohamed (2013). However, there is also a potential selectivity issue here that could explain the results and their observed temporal instability. That is, the types of drivers owning these vehicles may change over time and this result in different effects. For example, the 2013 to 2017 time period corresponds to a rapid increase in the vehicle market penetration of SUVs. It could be that the risk profiles of drivers who first join the SUV movement by buying an SUV differ from those who buy SUVs later. This would be reflected in a changing impact of the SUV indicator. A similar phenomenon has been empirically modeled in Winston et al. (2006) where they found that the safest drivers are attracted to vehicles with advanced safety features first, making the effectiveness of the safety features seem high initially, but less so as riskier drivers start purchasing these advanced safety-feature vehicles. Unfortunately, developing an econometric model to capture this selectivity issue and thus uncovers the true effect of the vehicle type would be extremely challenging as discussed in Mannering et al. (2020).

1.6.3 Roadway-related characteristics

The traffic signal control indicator variable was statistically significant in some models and exhibited unstable behavior over time. Interestingly, this variable had opposite effects between daytime and nighttime in 2015 models with the probability of severe injury being lower in the daytime and higher in the nighttime when traffic signals were present (see Tables 8 and 9).

The no-pedestrian-signal indicator variable was a statistically significant variable in two models for each time of day showing some stability in daytime models while it showed an unstable behavior in the nighttime models. For daytime, Tables 8 and 12 show that this variable was statistically significant in the 2015 and 2017 models with the average effect being a decrease in no visible injury and an increase in moderate and severe injuries. For nighttime, this variable was statistically significant in 2013 and 2015 with opposite average effects in moderate and severe injury severities between the two years (see Tables 5 and 9). Model estimation results show that the low-speed limit indicator variable (30 miles per hour and less) was statistically significant in six models: in daytime (2014, 2016 and 2017) and nighttime (2015, 2016 and 2017), showing a relatively stable behavior over the years and different times of day. Marginal effects in Table 14 show that the average effect of this variable was to decrease the likelihood of severe injury in all models and increase the likelihood of no visible injury in all models (except for 2017-nighttime model) confirming the findings of previous studies.

The multiple lanes indicator variable (3 lanes or more) was also found to have a consistent behavior between daytime and nighttime over the years. Marginal effects in Table 14 show that this variable was statistically significant in two nighttime models (2013 and 2015) and one daytime model (2017) with the average effects being an increase in the likelihoods of severe and moderate injury and a decrease in the likelihood of no visible injury. Tables 7 and 8 show that the wet pavement indicator variable was statistically significant in two consecutive years but different time of day. Results in Table 14 show that this variable had opposite effects between daytime and nighttime in the likelihoods of no visible injury being decreased in nighttime accidents (in 2014) and increased during daytime accidents (in 2015). Similar to the findings by Zamani et al. (2021), the road surface condition variables (both wet and dry) were only significant in some years with somewhat unstable behavior.

1.6.4 Crash-related characteristics

The vehicle-front impact indicator variable was found to be statistically significant in various years and times of day with the effects being somewhat stable. For daytime models, marginal effects in Table 14 show that the average effects of this variable decreased the likelihood of no visible injury in three consecutive years (2014, 2015, and 2016) and increased the likelihood of severe injury in those years, except for 2016, with relatively stable magnitudes. For the nighttime models, this variable was found to follow the daytime effects for two consecutive years (2016 and 2017). The severe injury outcomes resulting from vehicle frontal impact is expected due to the mass transfer from a moving vehicle to a pedestrian. Accidents occurring on the roadway but not at an intersection were found to be a statistically significant fixed-parameter variable in two nighttime models and one daytime model. In 2015, this variable was found to be statistically significant in both times of day with similar marginal effects, increasing the likelihood of severe injury and decreasing the likelihoods of no visible and moderate injuries. This is also expected because pedestrians are less likely to be expected at non-intersection locations on the roadway, which may increase the reaction time of drivers giving them less opportunity to reduce their speeds and take evasive action before a collision. This conclusion is somewhat supported by the findings of some previous studies that found that pedestrians suffered more severe injuries outside of crosswalks and less severe injuries in crosswalks (Mokhtarimousavi et al. 2020; Zamani et al. 2021)

1.6.5 Driver-related characteristics

In terms of drivers' gender, the male driver indicator variable shows relatively unstable behavior across different years and times of day. In the 2013 models, this variable indicator increased the likelihood of severe injury in the nighttime and increased the likelihood of moderate injury in the daytime.

Estimation results in Tables 4 and 12 show that the seatbelt indicator variable was statistically significant in two daytime models (2013 and 2017) with no significant variables in the nighttime models. Marginal effects in Table 14 show that the average effects of this variable were constant over the years in terms of direction (increasing the likelihood pedestrian severe injury) but lower in terms of the magnitude of the marginal effects in 2017. This seatbelt use variable finding suggests some compensating behavior among drivers. That is, as drivers feel safer (with seatbelt use) they may drive faster to keep their overall risk at roughly the same level, which supports the theory and findings of Peltzman (1975) and Winston et al. (2006).

The younger driver indicator variable (less than 25 years old) was found to be a statistically significant fixed variable in two years but at different times of day (see Tables 6 and 13). Marginal effects show that this variable had similar direction of the average effects in two different years and times of day (2014-daytime and 2017-nighttime) increasing the likelihoods of no visible and moderate injury and decreasing the likelihood of severe injury in both models, following the effects of the statistically significant random parameter of the same variable in 2015-nighttime model. This finding, however, contradicts the results of a previous study by Pour-Rouholamin and Zhou (2016) where adult drivers (16-24 years old) were found to increase the severe injury likelihood.

1.6.6 Pedestrian-related characteristics

Results in Tables 4 through 13 show that pedestrian age was found to be a statistically significant variable in several years in the two times of day with a relatively consistent finding in terms of impact on injury-severity probabilities. The younger pedestrian indicator variable (less than 25 years old), for example, was statistically significant in both times of day in three consecutive years (2013, 2014, and 2015) with the effects of severe injury being decreased in all models (see Table 14). In contrast, the older pedestrian indicator (more than 55 years old) was found to be a statistically significant variable in two nighttime models (2013 and 2016) and one daytime model (2016) with the severe injury effects being increased in these models (see Table 14). These findings confirm the results of the recent study by Zamani et al. (2021) where younger pedestrians (less than 31 years) were found to decrease the likelihood of severe injury while older pedestrians (over 50 years) were found to increase the likelihood of severe injury. Age (of both drivers and pedestrians) could be a proxy of several unobserved factors including differences in physical characteristics, reaction time, and risk-taking behavior (Mannering et al., 2016).

1.7 Predictive Comparisons (Prediction Simulations)

The findings of temporal instability over the years indicate that there has been a fundamental shift in the effect that explanatory variables have had on resulting injury severity probabilities. However, from a pragmatic and predictive perspective it would be interesting to determine what the aggregate effect of the observed shift in the influence of explanatory variables has had on injury severity probabilities. That is, if parameters determined from model estimates based on 2014 data were used to estimate injury probabilities of crashes observed in 2017, what would the differences be relative to actual observed 2017 injury probabilities? The answer to this question can be acquired via model prediction. However, prediction with random parameters models must be carefully done to account for parameter variance because it can be readily shown that simply using the mean of the random parameter for prediction is clearly incorrect and will result in biased forecasts (Xu et al., 2021). Several approaches to correctly forecast with random parameters models are available. For within-sample prediction (prediction with the same observations used to estimate the model) the simulated Bayesian approach described in Greene (2004) can be used to determine the parameters of individual observations, and these individual observation parameters can then be used to provide individual forecasts in response to changes in explanatory variables (see Washington et al., 2020 and Alnawmasi and Mannering, 2022 for applications of this approach). For out-of-sample prediction (prediction for a sample of observations that was not used for model estimation) the parameters of individual observations will not be transferable from one group of observations to another (since the observations will be different in the different data sets), so the Bayesian parameter estimation used for within-sample prediction cannot be used. The prediction undertaken in the current report will be out-of-sample because estimated parameters from one year/day/night model will be used to forecast with the observations from a different year/day/night data sample.[3] As a result, for out-of-sample predictions, the estimated parameters from the base sample (the sample of observations used to initially estimate the model), including the full distribution of random parameters (based on estimated means and variances), must be used and applied in a simulation method similar to that used in estimating the model (using either Halton or random draws to compute discrete outcome probabilities). Thus, out-of-sample prediction can be done by simulation, numerically integrating Equation 3 to compute individual crash injury-severity probabilities (in much the same way that Equation

[3] That is, year/day/night model combinations will be considered in-sample when their data is used for estimation and out-of-sample when their data is used for prediction using the parameters estimated in a different year/day/night combination.

3 is integrated for model estimation).[4] Hou et al. (2022) provide a detailed explanation, discussion, and empirical assessment of this technique.[5]

To begin the out-of-sample prediction, attention is first directed toward comparisons of day and night crashes. Presumably, the more severe pedestrian injury severities observed at night result largely from poorer visibility due to inadequate lighting relative to day conditions (although driver and pedestrian fatigue and other factors may also play a role). An interesting question would then be what would the nighttime severity distributions be if daytime parameter estimates were used to forecast them? The results of this out-of-sample simulation (using daytime models to predict nighttime injury severity given observed nighttime crash characteristics) is present in Table 15, which provides a summary of changes in pedestrian injury severity prediction means between day and night in all years from 2013 to 2017, with Figures 1 to 5 providing corresponding histograms of the distribution of probability differences for individual crash observations.[6]

The overall results of the prediction simulations shown in Table 15 indicate that the pedestrian injury severity differences between day and night are relatively stable in terms of direction, but with different magnitudes. The means of injury severity outcomes in Figure 1 show that using daytime data in 2013 to predict nighttime probabilities overestimates no visible and moderate injuries by 0.0260 and 0.0960, respectively, and underestimates severe injury by

[4] Some software packages offer this out-of-sample simulation for the random parameters logit model as part of a standard prediction routine, where the same simulation approach used for estimation is used to predict out-of-sample outcome probabilities. However, the simulation can be done without the use of estimation software using the approaches discussed and demonstrated in Hou et al. (2022).

[5] For additional information on this process, please see recent studies by Hou et al. (2021) and Xu et al. (2021) for a discussion of out-of-sample prediction using random parameters in the context of count-data models, and Alogaili and Mannering (2020), Islam et al. (2020) and Alnawmasi and Mannering (2021) for previous out-of-sample forecasts using injury severity models.

[6] The improved visibility resulting from daytime conditions could also result in a fundamental shift in the sample of crash-involved drivers, since the improved visibility may allow the safest drivers to take evasive maneuvers to not only reduce injury severities (Fountas and Anastasopoulos, 2018) but also possibly avoid the crash completely. This results in a potential sample selection issue as discussed in Mannering et al. (2020), where the risk-prone characteristics of crash-involved drivers will vary by time of day. Given this, some caution should be exercised in interpreting the simulation findings.

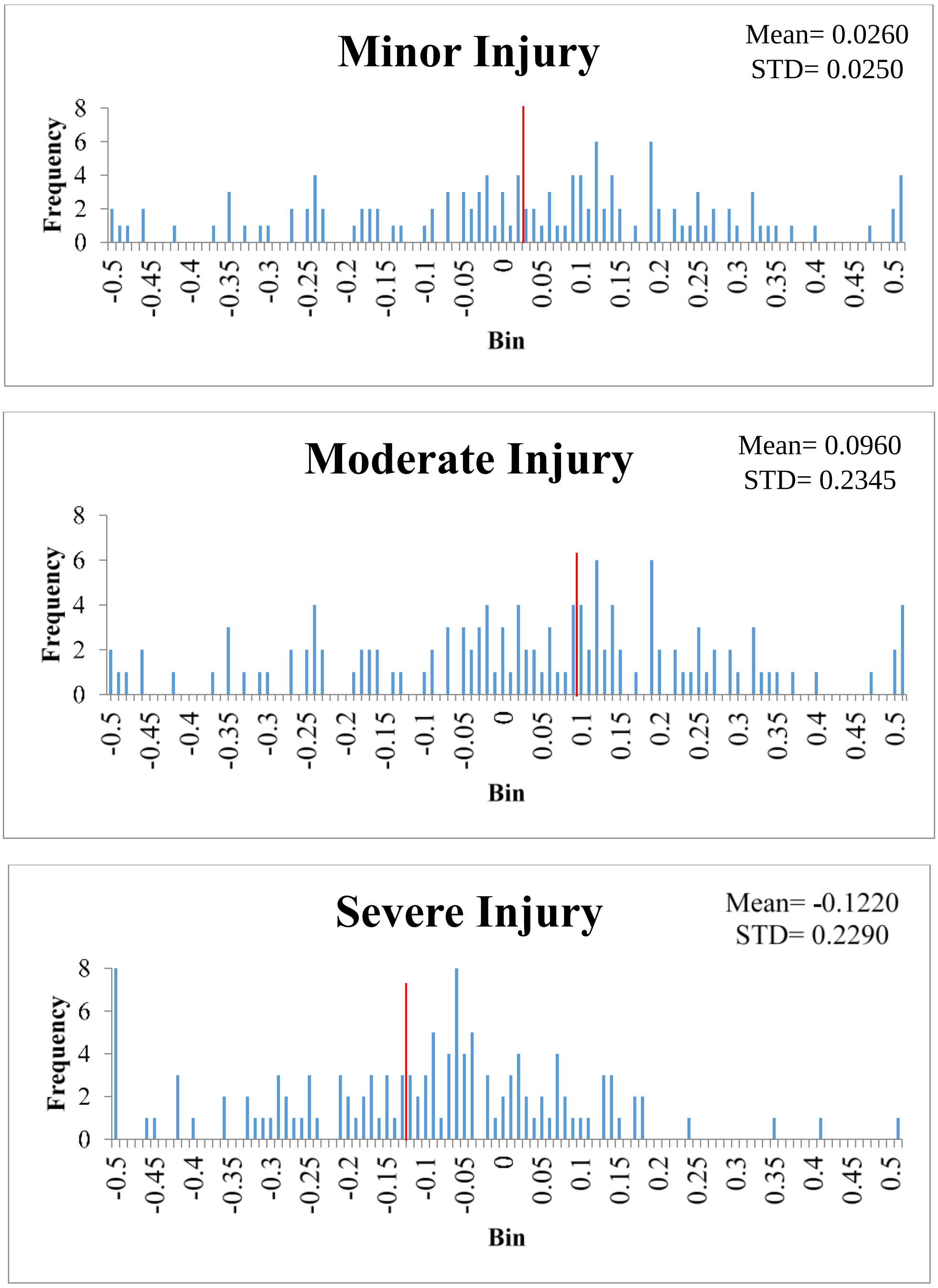

Figure 1. Difference between the pedestrian 2013-night estimated model predicted injury probabilities using 2013-day data and 2013-night "observed" probabilities.

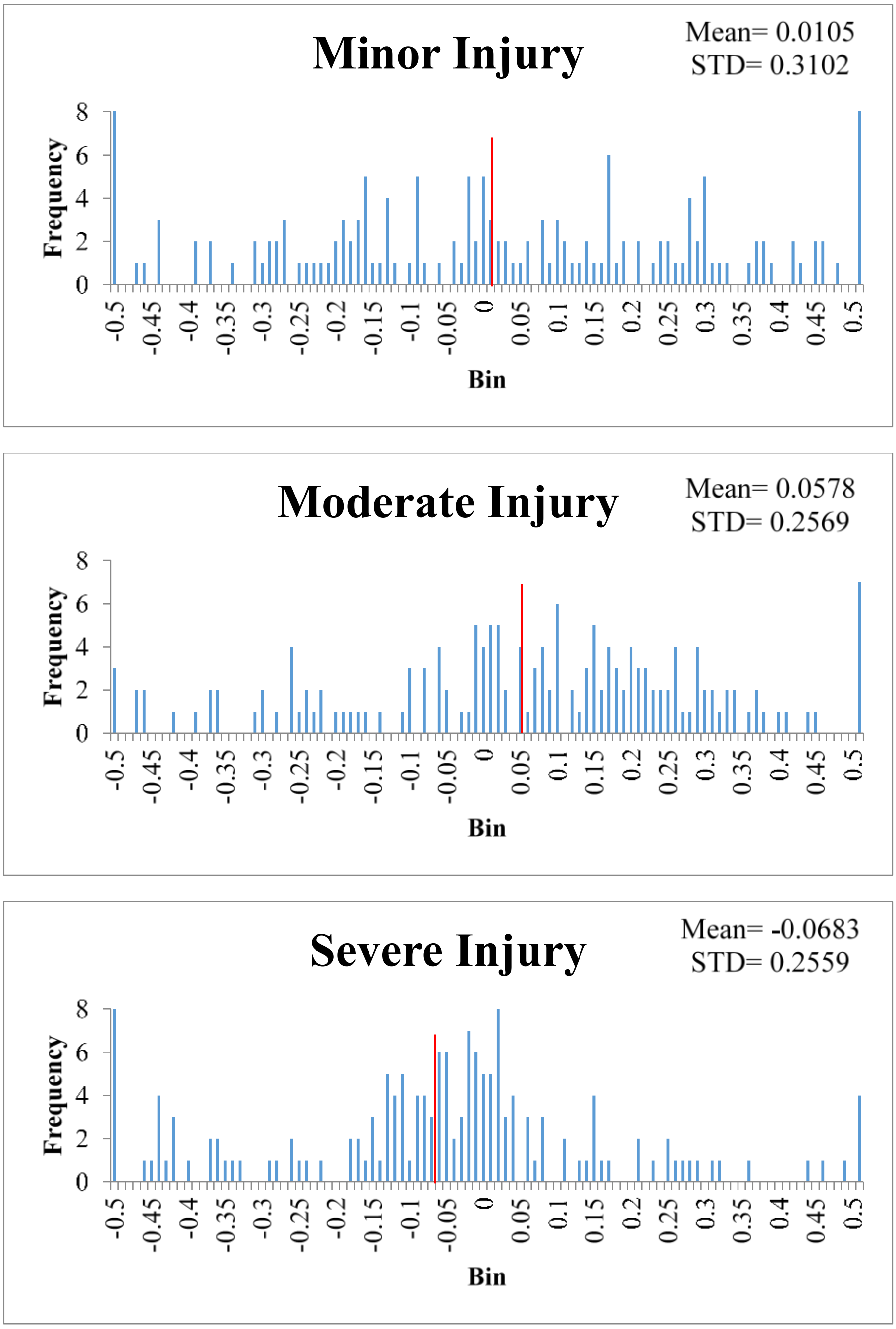

Figure 2. Difference between the pedestrian 2014-night estimated model predicted injury probabilities using 2014-day data and 2014-night "observed" probabilities.

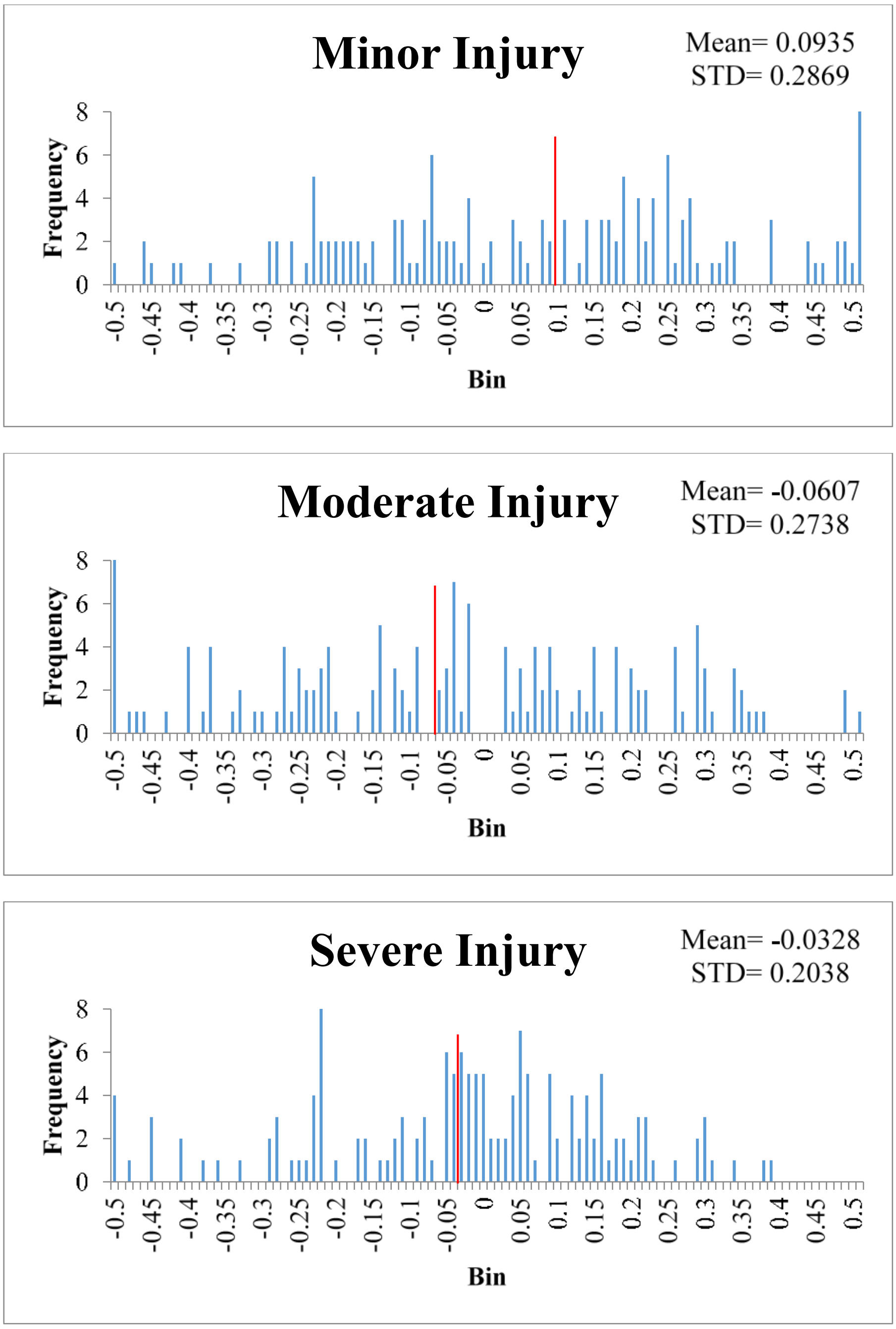

Figure 3. Difference between the pedestrian 2015-night estimated model predicted injury probabilities using 2015-day data and 2015-night "observed" probabilities.

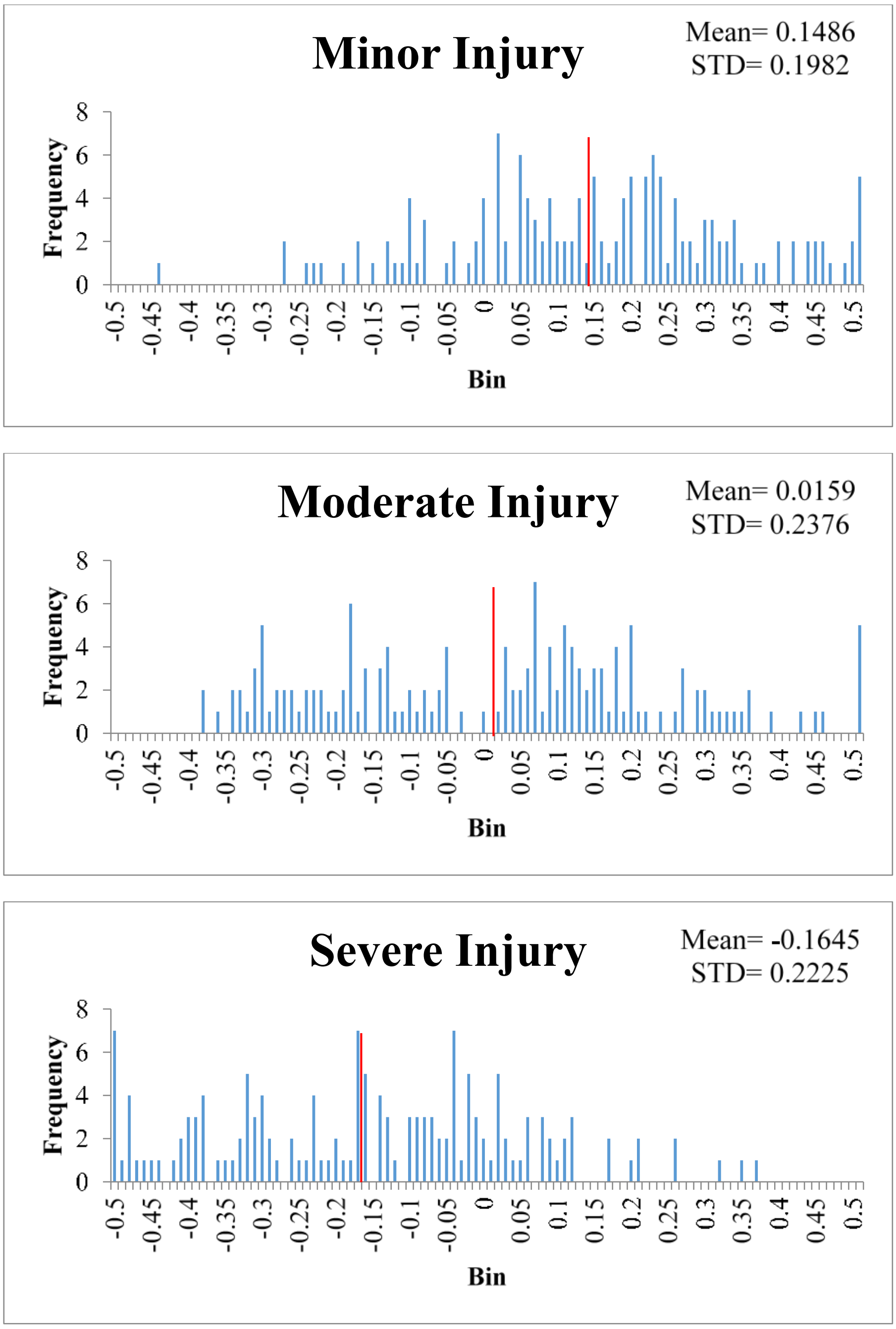

Figure 4. Difference between the pedestrian 2016-night estimated model predicted injury probabilities using 2016-day data and 2016-night "observed" probabilities.

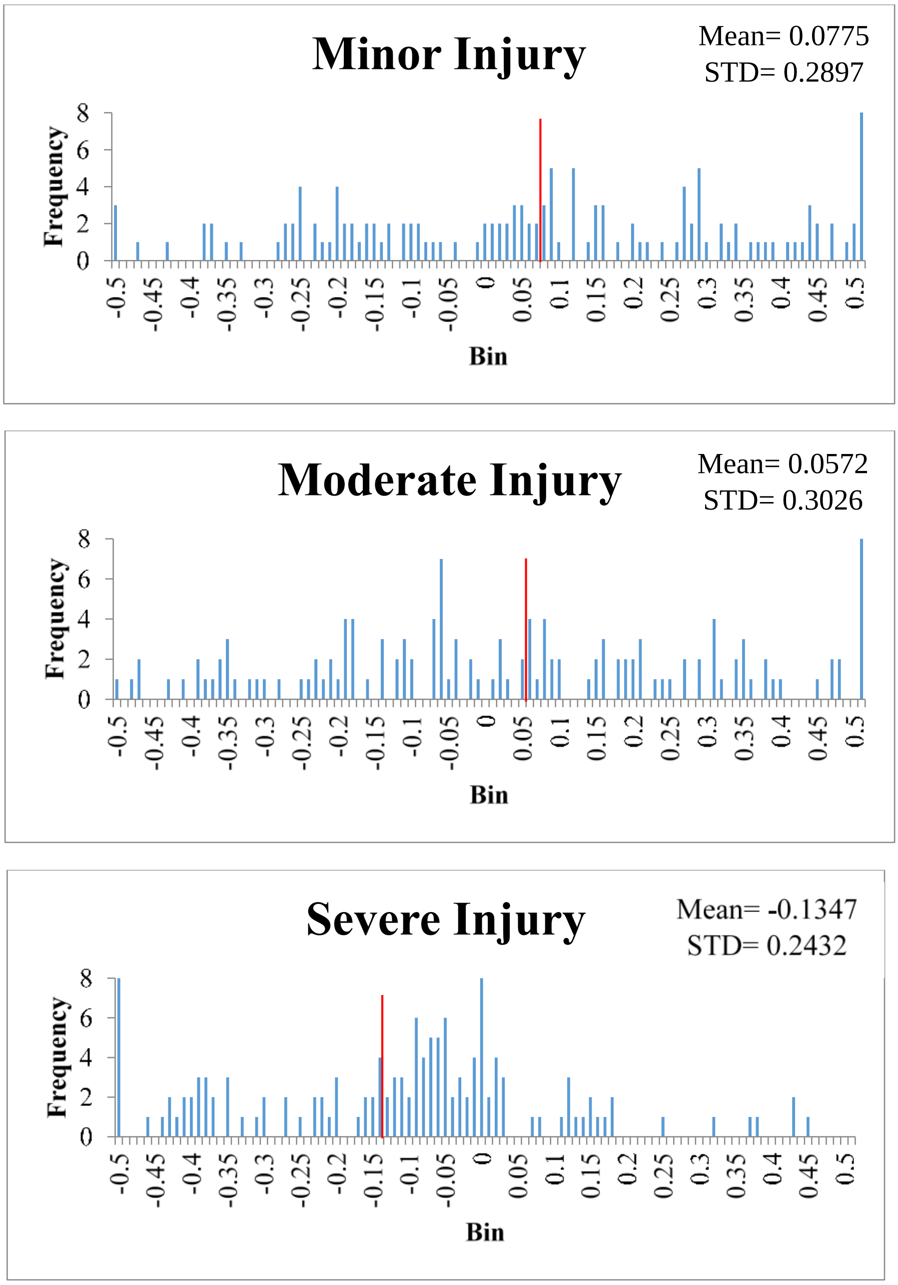

Figure 5. Difference between the pedestrian 2017-night estimated model predicted injury probabilities using 2017-day data and 2017-night "observed" probabilities.

0.1220. Meaning, that daytime estimated parameters would predict more no visible and moderate accidents but less severe accidents than observed. Table 15 shows that injury severity outcomes follow this pattern in all years overestimating no visible injury and moderate injury (except for 2015) and underestimating severe injury. Although the direction of the predictions is relatively stable over the years, magnitudes of the severe injury "underestimation" in the recent years (2016 and 2017) are much larger, which suggests an increasing separation between day and night injury severities. The findings in Table 15 are important because they show, in some sense, the upper limit of what can be achieved by improving lighting and pedestrian visibility or having pedestrian detection capabilities in vehicles, since the predictions (using daytime parameters to predict nighttime severities) give insights as to what could be achieved by essentially replicating daytime conditions. The findings suggest a rather dramatic effect because the 0.1645 and 0.1347 reduction in severe injury probabilities in 2016 and 2017 would result in thousands of fewer incapacitating injuries and fatalities across the nation.[7]

Table 15. Summary of changes in pedestrian injury severity prediction means between day and night by the year of the accident.

Year	No visible Injury	Moderate Injury	Severe Injury
2013	0.0260	0.0960	-0.1220
2014	0.0105	0.0578	-0.0683
2015	0.0935	-0.0607	-0.0328
2016	0.1486	0.0159	-0.1645
2017	0.0775	0.0572	-0.1347

To study the aggregate effect of the temporal instability in day and night crashes further, Table 16 provides the differences between day and night forecasted probabilities and day and night "observed" probabilities for all years (2013-2017) using 2013 to 2016 base years. As an

[7] Please note that some of results in Table 15 are based on year/day/night models that were estimated with relatively few observations. The use of larger data bases should provide precise probability estimates.

Table 16. Prediction results (base year model forecasted minus observed forecast year predicted probabilities) for day and night accidents for: no visible injuries (first number), moderate injury (in parentheses) and severe injury (in brackets) for day and night.

	Forecast Year							
	2014		2015		2016		2017	
Base Year	Day	Night	Day	Night	Day	Night	Day	Night
2013	-0.0679 (-0.0066) [0.0745]	-0.0207 (-0.0572) [0.0779]	-0.0746 (0.0373) [0.0373]	0.0044 (-0.0619) [0.0575]	-0.0643 (-0.0091) [0.0734	0.0561 (-0.0949) [0.0388]	-0.0172 (-0.0305) [0.0477]	0.0169 (-0.0326) [0.0157]
2014			0.0129 (0.0300) [-0.0429]	0.0960 (-0.0593) [-0.0367]	-0.0256 (0.0044) [0.0212]	0.1394 (-0.0698) [-0.0695]	0.0293 (-0.0024) [-0.0269]	0.0811 (-0.0180) [-0.0632]
2015					-0.0215 (-0.0314) [0.0529]	0.0356 (0.0061) [-0.0417]	-0.0091 (-0.0269) [0.0360]	0.0116 (0.0201) [-0.0317]
2016							0.0399 (-0.0038) [-0.0361]	-0.0572 (0.0312) [0.0260]

example of how to read Table 16, the table shows that if the 2013 nighttime model parameters were used to forecast 2017 nighttime injuries, no visible injuries would be overestimated (relative to observed injury distributions in 2017 nighttime) by 0.0169, moderate injuries would be underestimated by 0.0326, and severe injuries would be overestimated by 0.0157.

The results show that 2013 daytime and nighttime models overestimate all subsequent years' no visible injury while for nighttime, 2013 model overestimates subsequent years' no visible injury (except for 2014). For 2014 (as a base year), the daytime model overestimates no visible injury severity for all subsequent years (except for 2016) while for nighttime, it follows the predications of 2013 base year model with higher magnitudes.

For moderate injury severity outcomes, Table 16 shows (values in parentheses) that 2013 daytime model underestimates all subsequent years' moderate injury. Interestingly, moderate injury parameters are also underestimated in all nighttime models using 2013 model as a base year. However, when 2015 and 2016 nighttime models were used to predict subsequent years' parameters moderate injury outcomes were overestimated.

For severe injury severity outcomes, Table 16 shows (values in brackets) that using 2013 models (daytime and nighttime) overestimates severe injury of all subsequent years. In contrast, using 2014 models (both times of day) underestimates severe injury of all subsequent years (except for 2016 daytime model). Overall, the findings of the prediction simulations present evidence on the rather substantial aggregate effect of the temporal instability of pedestrian injury severities over the years.

1.8 Pedestrian Findings Summary

In the U.S., pedestrian fatalities have increased substantially in recent years and have been accounting for an increasing percentage of overall roadway-related fatalities. There are a multitude of reasons for why this may be happening. One possibility is that the safety advancements implemented in vehicles recently might be increasing drivers' risk-taking behavior as they compensate for the improved safety that such advancements provide (Winston et al., 2006), making pedestrians at a greater risk. There is also the potential issue of increased cell phone use by both drivers and pedestrians, resulting in potential distractions that may result in increased pedestrian-injury severities. Finally, changes in vehicle fleet composition (the trend toward larger vehicles such as sport utility vehicles, with greater mass and potentially less pedestrian-friendly frontal areas) may also be playing a role in increases in pedestrian injuries. To explore these possibilities, the current report uses data from single-pedestrian single-vehicle crashes in Kansas from 2013 and 2017 and estimates a series of random parameters logit models with heterogeneity in the means and variances to study the relationship between the severity of pedestrian accidents and time of day over the years.

Many statistically significant variables were found to affect the pedestrian injury severity probabilities in day and night conditions over the years studied. Some of these variables, such as wet pavement and traffic signal indicator variables, were found to have an opposite effect between day and night. Other variables were found have generally similar effects on both times of day including an indicator for accidents that occurred on roadway but not on intersection and the younger driver indicator variables. Variables that were found to increase the likelihood of severe injury or decrease the likelihood of no visible injury in multiple years included indicators for larger vehicle types (Pickups and SUVs), frontal vehicle impact, no-pedestrian sign, multiple lanes , and older pedestrians. Other variables were found to decrease the severe injury outcome such as indicators for younger drivers, younger pedestrians, and lower speed limits. Several models produced a statistically significant random parameters with heterogeneity in the means which capture unobserved heterogeneity present in the data.

In terms of temporal stability, the results show some differences in the behavior of specific indicator variables between day and night over the years. For example, indicators for younger drivers, older pedestrians, lower speed limits, multiple-lane roadways, and frontal vehicle impact were found to have a relatively stable behavior throughout the years in both times of day, while indicators for pickups, and SUVs, and no-pedestrian signs were found to be stable only during the day. While male driver and warmer-weather accident indicator variables were temporally unstable in both times of day, indicators for pickups and no-pedestrian sign were temporally unstable in nighttime models.

Predictive comparison results show some evidence of temporal stability between day and night predictions where using daytime parameter estimates to predict nighttime probabilities overestimates no visible injury and underestimates severe injury in all studied years. However, when predicting the probabilities of each time of day in several subsequent years, there was a variation in the prediction simulation results that show clear evidence of temporal instability throughout the years.

The results of this study also show a clear evidence that factors that affect pedestrians' injury vary significantly between day and night conditions but suggest that severe pedestrian injuries could be reduced by replicating day conditions at night through improved illumination, pedestrian detection systems in vehicles, and methods to mitigate potential driver fatigue at night.

Policies along these lines could potentially result, annually, in thousands of fewer incapacitating injuries and fatalities across the nation.

More importantly, the results underscore the need to account for unobserved heterogeneity in the analysis of crash data. The strong statistical significance of random parameters is a clear indication that unobserved effects relating to missing attitudinal and behavioral data are playing an important role. Because current highway safety practice (AASHTO, 2010) ignores these unobserved effects, erroneous conclusion and inferences may be drawn. The findings in this part of the report clearly underscore the need to consider unobserved effects in statistical estimation, through the use of mixing distributions, to account for missing data which includes, among other concerns, the safety-related attitudes of road users.

Part II

An assessment of no-injury and injury crash occurrences: Bivariate versus univariate models with alternate random parameters structures

2.1 Introduction and literature review

The Highway Safety Manual (AASHTO, 2010) has become the mainstay of safety analysis in the United States. However, the current version of the manual uses traditional statistical methods with no way of capturing the effect of motorist attitudes and unobserved behavioral elements of highway safety. This can potentially lead to biased estimates of model parameters as well as inaccurate inferences on the effectiveness on the safety features and/or policies that analysts wish to explore. Moreover, the structure of the statistical models used in the Highway Safety Manual (commonly referred to by the acronym HSM) often overlooks potentially important correlations among the frequencies of injury severities on roadway segments. This can lead to further inaccuracies in policy assessment. The intent of this portion of the report is to demonstrate an additional application of mixing distributions, in this case to both the frequency of crashes as well as their resulting injury severity (the preceding portion of this report focused exclusively on injury severity in its application of mixing distributions).

This portion of the report focusses on identifying the factors affecting the frequencies of non-injury and injury crashes on freeways. Poisson models are one of the early models used in crash frequency modeling. An inherent restriction of mean and variance for dependent variable to be equal in Poisson models results in invalid results when this condition is violated. Negative binomial models relax the condition of equal mean and variances and can effectively take care of this over-dispersion problem. Since the crashes are rare events and the crash data commonly consists of excess zeroes, both Poisson and negative binomial models result in biased results when used for datasets with excess zeroes. First approach in the literature to overcome the issue involves zero-inflated Poisson and zero-inflated negative binomial models. These models assume that the road locations exist in one of the dual states (zero state and non-zero state) in relation to the crashes (Shankar et al., 1997; Lee and Mannering, 2002; Lord et al., 2007; Malyshkina and Mannering, 2010). However, these models prioritize the statistical fit instead of explaining the crash occurrence process. A drastic improvement in computation power in the recent decades helped the safety researchers to conduct advanced studies with larger datasets and more complex models. Malyshkina et al. (2009) and Malyshkina and Mannering (2010) proposed a two state Markov switching model to allow road locations to change in time from one state to another.

Second approach to account for excess zeroes in the crash dataset is by considering a new distribution which can address lower counts and combining it with the parent Poisson and Negative binomial distributions. Recently developed models with such distributions are Poisson-Lindley, negative binomial-Lindley and negative binomial generalized exponential. Unobserved heterogeneity due to other sources is another issue in the count models. Unobserved heterogeneity can be classified into structured and unstructured. Structured heterogeneity can be due to temporal correlations as commonly seen in panel data where multiple observations are recorded over time. Random effects Poisson and Negative binomial models are commonly used to account for structured heterogeneity. The unstructured heterogeneity can be due to uncertainty in covariates and omitted variables. Ignoring such heterogeneity can lead to bias in parameter estimates and results in incorrect inferences from the model. Random parameters models address the heterogeneity by assuming a distribution for parameters and allowing the parameters to vary across observations. Table 17 summarizes the studies which used random parameters models for crash

Table 17. Summary of empirical studies that used univariate random parameters count models in accident research.

Study	Abbreviation(s) of model(s) considered	Outcome response count variable	Major findings/contributions from methodological standpoint
Anastasopoulos and Mannering (2009)	Fixed parameters negative binomial model and random parameters negative binomial model.	5 years motor vehicle crash frequency.	• Random/fixed parameters negative binomial are statistically better than the random/fixed parameters Poisson models. • Random parameters negative binomial model is statistically superior to fixed parameters negative binomial model at more than 99.99% confidence level.
Venkataraman et al. (2011)	Fixed parameters negative binomial model and random parameters negative binomial model.	Interstate crash frequency.	• Random parameters negative binomial model significantly improved the likelihood value and statistically better compared to the fixed parameters negative binomial model.
Ukkusuri et al. (2011)	Random parameters negative binomial model	Severe and fatal pedestrian-vehicle crash frequency.	• Random parameters negative binomial model has significant improvement in the likelihood value compared to restricted-at-zero model.
Venkataraman et al. (2013)	Fixed parameters negative binomial model and random parameters negative binomial model.	Interstate crash frequencies by crash severity, number of vehicles involved, collision and location type.	• Except for head on collisions, random parameters negative binomial model models have statistically better fit than their corresponding fixed parameters negative binomial model.
Mohammadi et al. (2014)	Random parameters negative binomial model and Random effects negative binomial model.	Interstate highway crash frequency.	• Random parameters negative binomial model fits data better but the random effects negative binomial model has slightly better predictive power because of fewer parameters in the model.
Coruh et al. (2015)	Correlated Random parameters Poisson model, uncorrelated random parameters Poisson model, correlated random parameters negative binomial model and uncorrelated random parameters negative binomial model.	Monthly crash frequency on highways across Turkey.	• Random parameters negative binomial models are statistically superior to the random parameters Poisson models. • Correlated random parameters models outperformed their corresponding uncorrelated models.
Caliendo et al. (2015)	Fixed parameters negative binomial model and random parameters negative binomial model.	Severe crash frequency in tunnels.	• Random parameters negative binomial model could not outperform the fixed parameters negative binomial model for tunnel crash data.

Naznin et al. (2016)	Random parameters negative binomial model and Random effects negative binomial model.	Tram involved injury crash frequency.	• Random effects negative binomial model is statistically better than the fixed parameters negative binomial model in terms of data fit.
Kamla et al. (2016)	Fixed parameters negative binomial model and random parameters negative binomial model.	11-year number of roundabout crash counts.	• Random parameters negative binomial model provided better data fit than fixed parameters negative binomial model.
Rusli et al. (2017)	Random parameters negative binomial model.	Monthly single vehicle crash counts	
Rusli et al. (2018)	Random parameters negative binomial model, random parameters negative binomial-Lindley model and random parameters negative binomial – generalized exponential model.	Multivehicle crash counts on mountainous highways.	• Random parameters negative binomial-Lindley outperformed random parameters negative binomial model and random parameters negative binomial – generalized exponential model in terms of prediction ability and model fit.
Shaon et al. (2018)	Negative binomial – Lindley, and random parameters negative binomial model, and random parameters negative binomial-Lindley model.	Crash frequency on rural interstate sections and two-lane rural roadway sections.	• Random parameters negative binomial-Lindley model performed statistically better fit compared to negative binomial – Lindley and random parameters negative binomial models • Random parameters negative binomial-Lindley model offered better understanding about the effects of potential underlying factors.

frequencies. Anastasopoulos and Mannering (2009), Venkataraman et al. (2011), Venkataraman et al. (2013), and Kamla et al. (2016) found that random parameters models are statistically superior to the fixed parameters in terms of data fit. However, Caliendo et al. (2015) found that random parameters models could not perform better than the fixed parameters for tunnel crash frequency data.

Interactions between the factors affecting the crashes could arise due to unknown sources and ignoring such interactions may result in inaccurate inferences. Correlated random parameters models can effectively capture such interactions and may provide better data fit than the uncorrelated random parameters models. Table 18 summarizes the existing correlated random parameters models used for modeling crash frequency modeling. Studies by Coruh et al. (2015), Hou et al. (2018), Caliendo et al. (2019), Huo et al. (2020) and Tang et al. (2020) found that correlated random parameters models provide statistically better fit than the uncorrelated models. However, Venkataraman et al. (2011) and Saeed et al. (2019) found that correlated random parameters models couldn't perform better than the uncorrelated random parameters models. Table 19 summarizes the random parameters models with heterogeneity in means and variances.

Instead of modeling the total crashes including no-injury and injury crashes, developing separate models provide more insights about the factors affecting different injury crashes. Since the factors affecting different injury crashes are correlated, estimating separate models may lead to incorrect inferences. Bivariate/multivariate models address this issue by modeling different injury crashes simultaneously. Table 20 summarizes the existing studies which used bivariate and multivariate models. Dong et al. (2014), Dong et al. (2017), Barua et al. (2015), Chen et al. (2017), Liu et al. (2018), Bhownik et al. (2019) and Wang et al. (2020) found that multivariate models are statistically better than the univariate models in terms of data fit.

Table 18. Summary of empirical studies that used correlated random parameters count models in accident research.

Study	Abbreviation(s) of model(s) considered	Outcome response count variable	Major findings/contributions from methodological standpoint
Venkataraman et al. (2011)	Fixed parameters negative binomial model, Uncorrelated Random parameters negative binomial model and Correlated Random parameters negative binomial model.	Total annual crash frequency on interstate segments.	• Random parameters negative binomial models provided a better fit for observed crash frequency than fixed parameters negative binomial model. • Uncorrelated random parameters negative binomial yielded the best likelihood when compared to Correlated random parameters negative binomial models.
Coruh et al. (2015)	Fixed parameters negative binomial model, Uncorrelated Random parameters negative binomial model and Correlated Random parameters negative binomial model.	Monthly crash frequency on highways across Turkey.	• The statistical superiority (high to low) of models in terms of data fit is correlated random parameters negative binomial model, uncorrelated random parameters negative binomial model and fixed parameters negative binomial model. • Random parameters models outperform very well in terms predicting monthly crash frequencies.
Hou et al. (2018)	Random effects negative binomial model, uncorrelated random parameters negative binomial model and correlated random parameters negative binomial model.	Annual crash frequency in tunnels.	• Uncorrelated random parameters negative binomial model provided better data fit compared to random effects negative binomial model. • Correlated random parameters negative binomial model outperformed uncorrelated random parameters negative binomial model and provided better goodness-of-fit and more insights on factors affecting the tunnel safety.
Caliendo et al. (2019)	Random effects Poisson model, uncorrelated random parameters Poisson model and correlated random parameters Poisson models.	Annual crash frequency in motorway tunnels.	• Correlated random parameters Poisson model provided a better data fit when compared to uncorrelated random parameters Poisson model and random effects Poisson model.
Saeed et al. (2019)	Fixed parameters negative binomial model, uncorrelated random parameters negative binomial model and correlated random parameters negative binomial model.	3-years injury crash and no-injury crash frequency on multilane highways.	• For both no-injury and injury crash frequencies, uncorrelated random parameters negative binomial models outperform their Fixed parameters negative binomial counterpart models. • Interestingly, correlated random parameters negative binomial models did not provide statistically better data fit than the corresponding uncorrelated random parameters negative binomial models.
Tang et al. (2020)	Fixed parameters negative binomial model, Fixed Parameters Negative Binomial-Lindley model, Uncorrelated random parameters negative binomial- Lindley model and correlated random parameters negative binomial-Lindley model.	Annual crash frequency in tunnels.	• The statistical superiority (high to low) of models in terms of data fit is correlated random parameters negative binomial-Lindley model, uncorrelated random parameters negative binomial- Lindley model Fixed Parameters Negative Binomial-Lindley model and Fixed parameters negative binomial model.

Table 19. Summary of empirical studies that used heterogeneity-in-means count models in accident research.

Study	Abbreviation(s) of model(s) considered	Outcome response count variable	Major findings/contributions from methodological standpoint
Venkataraman et al. (2014)	Heterogeneity-in-means random parameter negative binomial model and fixed parameters negative binomial.	Interchange and non-interchange crash frequencies.	• Heterogeneity-in-means random parameter negative binomial model is statistically superior to the fixed parameters negative binomial model in terms of data fit.
Huo et. al. (2020)	Random parameters negative binomial model and correlated random parameter negative binomial with heterogeneity in means models.	4-years crash counts on mountainous freeways.	• correlated random parameter negative binomial with heterogeneity in means model provided statistically better fit than the Random parameters negative binomial model at 99% level of confidence.
Huo et al. (2020)	Heterogeneity-in-means and variances random parameter negative binomial model, Random parameters negative binomial model and negative binomial model.	Multilane mountainous freeway crash frequency.	• Statistical superiority (high to low) of models in terms of data fit is Heterogeneity-in-means and variances random parameter negative binomial, random parameter negative binomial model, and negative binomial model. • Heterogeneity-in-means and variances random parameter negative binomial model and Random parameters negative binomial model are less precise than the negative binomial models when applied to out-of-sample data.

Table 20. Summary of empirical studies that used bivariate and multivariate count models in accident research.

Study	Abbreviation(s) of model(s) considered	Outcome response count variable	Major findings/contributions from methodological standpoint
Dong et al. (2014)	Multivariate negative binomial model, multivariate zero inflated negative binomial model and multivariate random parameters zero inflated negative binomial model.	Car-only, car-truck and truck only crash frequencies at urban signalized intersections.	• The statistical superiority (high to low) of models in terms of data fit is multivariate random parameter zero inflated negative binomial model, multivariate zero inflated negative binomial and multivariate negative binomial model. • Multivariate random parameter zero inflated negative binomial model predicts better than the multivariate zero inflated negative binomial model.
Barua et al. (2016)	Multivariate Random Parameters model with Heterogenous effects, Multivariate Random Parameters model with Spatial Heterogeneity, Multivariate Random Parameters model with Heterogenous effects and Spatial Heterogeneity and Univariate Random Parameters model with Heterogenous effects and Spatial Heterogeneity.	No-injury and Injury crash counts.	• All three Multivariate models are comparable, and no model can be preferred over other. • Multivariate spatial models outperformed separate univariate spatial models in terms of goodness-of-fit.
Dong et al. (2017)	Multivariate zero inflated Poisson model, multivariate random parameter zero inflated Poisson model, multivariate random parameter negative binomial model, multivariate random parameter zero inflated negative binomial model.	Property damage only, possible injury, non-disabling, disabling and fatal crash frequencies at urban signalized intersections.	• multivariate random parameter zero inflated Poisson and multivariate random parameter zero inflated negative binomial models performed better than the multivariate random parameter negative binomial model in terms of data fit. • Multivariate random parameter zero inflated Poisson model provided better data fit in comparison to multivariate random parameter zero inflated negative binomial model.
Chen et al. (2017)	Five separate multivariate random parameters negative binomial models for segments with poor, fair, fair to good, good, and excellent pavement condition.	Fatal, injury and no-injury crash counts.	• Likelihood ratio tests suggests that estimating different models based on pavement condition is better than a model for whole data.

Liu et al. (2018)	Multivariate random parameter zero inflated negative binomial model, Multivariate Poisson Log-Normal model, random parameters zero inflated Poisson model, random parameters zero inflated negative binomial models, multivariate zero inflated Poisson, multivariate zero inflated negative binomial model and multivariate zero inflated Poisson.	Annual sideswipe crash counts, annual rear end crash counts and annual other crashes.	• Compared to all other models, Multivariate random parameter zero inflated negative binomial model is statistically superior. • Multivariate zero inflated models are statistically superior to multivariate Poisson/negative binomial models. • Multivariate zero inflated Poisson/negative binomial models are more complex but fit the data better. • Multivariate models could estimate parameters more accurately than univariate models. • Multivariate Poisson Log-Normal model performed worse than all other multivariate zero inflated models.
Bhowmik et al. (2019)	Negative binomial, random parameters multivariate negative binomial, independent panel negative binomial model and mixed panel negative binomial model.	Rear end, angular, sideswipe, all single vehicle, other multiple vehicle, and non-motorized crash counts.	• Independent panel negative binomial and mixed panel negative binomial models are statistically superior to the negative binomial and random parameters multivariate negative binomial models respectively. • Mixed panel negative binomial model provided better fit than the independent panel negative binomial model. • Similarly, random parameters multivariate negative binomial model performed better than the negative binomial model.
Wang et al. (2020)	Univariate negative binomial model, univariate Poisson model, multivariate zero inflated negative binomial model, multivariate zero inflated Poisson, random parameter multivariate negative binomial and random parameter multivariate Poisson models.	Rear-end, bumping guard rail, non-casualty, casualty, and other crashes.	• Zero inflated models are not suitable for the data used in the study. • random parameters multivariate negative binomial model is statistically superior to the random parameters multivariate Poisson model only for rear end, non-casualty, casualty, and other crashes.

2.2 Methodology

Crash frequency data represent the number of crashes occurred in a given period (one year) and are usually modeled using count data modeling techniques. Poisson models and its derivatives including zero-inflated models and negative binomial models are commonly used in count data modeling. According to the Poisson model, the probability $P(n_i)$ of road segment i having n_i crashes in a given period of time is denoted by

$$P(n_i) = \frac{EXP(-\lambda_i)\lambda_i^{n_i}}{n_i!} \tag{7}$$

Where λ_i is the expected number of crashes ($E(n_i)$) for road segment i. According to the Poisson regression, λ_i is represented as a function of explanatory variables using a log-linear function:

$$\lambda_i = EXP(\boldsymbol{\beta}\mathbf{X}_i) \tag{8}$$

Where $\mathbf{X}_i$ is a vector of explanatory variables and $\boldsymbol{\beta}$ is a vector of estimable parameters. The major limitation for Poisson model is that the mean and variance are equal ($E[n_i] = Var[n_i]$) for Poisson distribution. This condition may not hold true in all cases especially for over or under dispersed data. Using Poisson model for over/under dispersed data may results in the inaccurate estimation of standard errors and therefore incorrect inferences could be drawn. Negative binomial model addresses this issue effectively and equation (8) is rewritten as

$$\lambda_i = EXP(\boldsymbol{\beta}\mathbf{X}_i + \varepsilon_i) \tag{9}$$

Where, $EXP(\varepsilon_i)$ is a gamma-distributed error term with mean 1 and variance α. The negative binomial probability density function is:

$$P(n_i) = \left(\frac{1/\alpha}{(1/\alpha)+\lambda_i}\right)^{1/\alpha} \frac{\Gamma[(1/\alpha)+n_i]}{\Gamma(1/\alpha)n_i!} \left(\frac{\lambda_i}{(1/\alpha)+\lambda_i}\right)^{n_i} \tag{10}$$

Where $\Gamma(.)$ is a gamma function. As α approaches zero, the negative binomial model converges to Poisson model. To account for heterogeneity due to unobserved factors, random parameters are generally used for Poisson/negative binomial models. In random parameters count model, the estimable parameters are expressed as

$$\beta_i = \beta + \varphi_i \tag{11}$$

Where φ_i is a randomly distributed term such as normally, triangular, or uniformly distributed. For a random parameters model, equation (2) and equation (3) changes to $\lambda_i|\varphi_i = EXP(\boldsymbol{\beta}\mathbf{X}_i)$ and $\lambda_i|\varphi_i = EXP(\boldsymbol{\beta}\mathbf{X}_i + \varepsilon_i)$ respectively. The log-likelihood function can be written as

$$LL = \sum_{\forall i} \ln \int_{\varphi_i} g(\varphi_i) P(n_i|\varphi_i)\, d\varphi_i \tag{12}$$

Where g(.) is the probability density function of the φ_i. Most popular simulation based maximum likelihood method can be exercised and 1000 Halton draws are used in this report. Another more sophisticated approach to capture the unobserved heterogeneity is by allowing the mean and variances of random parameters vary by explanatory variables. Heterogeneity in the means of random parameters is introduced by specifying the $\boldsymbol{\beta}_i$ as

$$\boldsymbol{\beta}_i = \beta + \boldsymbol{\Theta}\mathbf{Z}_i + \xi_i \tag{13}$$

where β is the mean parameter estimate across all observations, $\mathbf{Z}_i$ is a vector of explanatory variables for an individual "i" which effects the mean of $\boldsymbol{\beta}_i$, $\boldsymbol{\Theta}$ is a vector of estimable parameters and ξ_i is a vector of randomly distributed terms. Heterogeneity in the variances of random parameters can also be introduced as (Washington et al., 2020)

$$\boldsymbol{\beta}_i = \beta + \boldsymbol{\Theta}\mathbf{Z}_i + \sigma_i EXP(\boldsymbol{\psi}_i \mathbf{W}_i)\xi_i \tag{14}$$

$\mathbf{W}_i$ is a vector of explanatory variables which captures the heterogeneity in standard deviation σ_i of random parameters and $\boldsymbol{\Psi}_i$ is the corresponding parameter vector.

Instead of modeling the non-injury and injury crash counts separately, a bivariate random parameters negative binomial model is used to model them simultaneously. According to the Winkelmann (2008), the expected number of crashes in bivariate negative binomial model can be expressed as

$$\lambda_{ik} = EXP(\boldsymbol{\beta}_k \mathbf{X}_{ik} + \varepsilon_{ik}) \tag{15}$$

Where k represents the non-injury and injury severity levels, λ_{ik} is the expected number of k^{th} level severity crashes in i^{th} road segment, $\mathbf{X}_{ik}$ is the vector of independent variables, $\boldsymbol{\beta}_k$ is the vector of coefficients, and EXP(ε_{ik}) is gamma distributed error term with mean 1 and variance α. If α is not significantly different from 0, Poisson model is preferred over negative binomial model.

Contrast to univariate negative binomial model, the error term ε_{ik} is multivariate error term which is based on unstructured correlated covariance matrix:

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho_{12} \\ \sigma_2\sigma_1\rho_{21} & \sigma_2^2 \end{bmatrix} \tag{16}$$

Where σ_1, σ_2 are standard deviations of error terms for non-injury and injury crashes. ρ_{12} is the correlation between the two error terms.

As discussed in the methodology section for univariate model for total crashes, restricting the effect of explanatory variables to be fixed for all observations may not appropriate and may results incorrect inferences. Instead, parameters are random and vary for different observations according to

$$\beta_n = \beta + \omega_n \tag{17}$$

where ω_n is a randomly distributed term (mixing distribution) and is typically assumed to be normally distributed with mean 0 and variance σ^2 (Washington et al., 2020)

2.3 Data

Data for this portion of the study is collected from 5 states including Georgia, Hawaii, Minnesota, Ohio, and Virginia. A total road stretch of 164.8 miles is divided into 728 homogenous sections based on their geometric characteristics and the resulting sections vary in length from 80 feet to 1.27 miles. Too smaller sections with length less than 0.1 miles are removed for estimation purposes and it resulted in a total of 478 sections. The number of injury and non-injury crashes occurred in these sections are observed for different time frames ranging from 1 year to 5 years. Variables available for estimation process are section related variables such as segment length, average annual daily traffic, number of lanes per section, inside and outside shoulder widths, proportion of section with inside and outside rumble strips, proportion of section with outside and inside barrier, offset to inside and outside barrier, distance to the nearest downstream exit ramp and upstream entrance ramp, average annual daily traffic of downstream exit ramp and upstream entrance ramp. All the available variables are tested in the modeling process and the descriptive statistics of statistically significant variables in the estimated models are presented in Table 21.

2.4 Estimation Results

The estimation results of heterogeneity in means random parameters negative binomial model[8] for total crashes obtained after extensive specification testing are presented in Table 22. Statistically significant dispersion parameter in Table 22 suggests that considering negative binomial model is more appropriate over Poisson regression model. The mean parameter for shoulder width is negative suggesting that an increase in the shoulder width more likely reduces

[8] Heterogeneity in the variances of random parameters is tested extensively but none of the variables are statistically significant. Similarly, correlations between the random parameters are turned out to be statistically insignificant.

Table 21. Descriptive statistics of variables used for modeling the crash frequencies on interstates.

Variable description	Mean	Standard deviation
Shoulder width in feet	9.67	3.46
Georgia state indicator (1 if the section is in Georgia, 0 otherwise)	0.09	0.28
Hawaii state indicator (1 if the section is in Hawaii, 0 otherwise)	0.07	0.26
Ohio state indicator (1 if the section is in Ohio, 0 otherwise)	0.26	0.44
Logarithm of section length in miles	-1.39	0.59
Logarithm of number of years the section is active	1.55	0.20
Logarithm of lane AADT	4.04	0.43
Proportion of section with outside rumble strip	0.28	0.43
Width of inside center barrier in feet	2.65	3.44
Distance from the nearest downstream exit ramp in miles	-0.69	1.14

the frequency of crashes. However, two variables, Georgia and Hawaii state indicators found to have different effects on the mean of the random parameter for shoulder width. A positive parameter estimate for Georgia state indicator suggests that an increase in shoulder width has lesser impact on reducing the crash frequency compared to other states (Minnesota, Ohio, and Virginia). Conversely, an increase in shoulder width has higher impact on reducing crash frequency in Hawaii when compared to the other states.

The parameter for logarithm of lane AADT is positive, suggesting that an increase in the lane AADT may more likely result in the higher crashes on interstates. A positive parameter estimate of Ohio state indicator suggests that the number of crashes are more likely to be higher in Ohio when compared to other states. Rumble strips are found to improve the road safety in the

Table 22. Estimation results of random parameters negative binomial model with heterogeneity in mean for total crash counts on interstates.

Variables	Parameter Estimates	t-statistic	Marginal effect
Random parameter (normally distributed)			
Constant	-3.535	-8.50	
(Standard deviation of parameter distribution)	(0.52)	(21.18)	
Heterogeneity in the mean of random parameter			
Constant: Georgia state indicator (1 if the section is in Georgia, 0 otherwise)	-1.175	-8.21	
Constant: Hawaii state indicator (1 if the section is in Hawaii, 0 otherwise)	-0.705	-2.90	
Random parameter (normally distributed)			
Shoulder width in feet	-0.04	-4.24	-0.90
(Standard deviation of parameter distribution)	(0.037)	(15.79)	
Heterogeneity in the mean of random parameter			
Shoulder width in feet: Georgia state indicator (1 if the section is in Georgia, 0 otherwise)	0.038	2.34	
Shoulder width in feet: Hawaii state indicator (1 if the section is in Hawaii, 0 otherwise)	-0.088	-3.13	
Fixed parameters			
Logarithm of section length in miles	1.084	25.57	10.80
Logarithm of number of years the section is active	1.042	7.30	6.42
Logarithm of lane average annual daily traffic	1.674	25.12	1.87
Ohio state indicator (1 if the section is in Ohio, 0 otherwise)	0.666	8.09	23.64
Proportion of section with outside rumble strip	-0.372	-4.52	-0.11
Width of inside center barrier in feet	0.019	2.52	0.55
Logarithm of distance from the nearest downstream exit ramp in miles	0.106	4.94	4.1
Dispersion parameter	5.397	11.32	
Number of observations		478	
Log-likelihood at convergence		-1894.14	
Constants only log-likelihood		-2106.54	

existing safety literature and the negative parameter estimate for proportion of section with outside rumble strip suggest that an increase in the percentage of rumble strips in a section may more likely reduce the number of crashes. Conversely, the number of crashes is more likely to increase with an increase in the width of center barrier. Surprisingly, an increase in the distance from downstream exit ramp more likely to increase the crash frequency.

The marginal effect of each variable on crash frequency is presented in the Table 22. It is important to note that increments used for computing marginal effects are different for each variable. Marginal effect of shoulder width is -0.9 implying that a 1-foot increase in the shoulder width may result in an average decrease of 0.9 crashes in a section. A 0.1-mile increase in the section length may result in 10.80 more crashes on average. A 1-year increase in the observation period may result in 6.42 more crashes on average. Similarly, a 1,000 increase in the lane AADT may result in 1.87 more crashes. A section in Ohio is more likely to have 23.64 more crashes than other states. A 1 percent or 0.01 increase in the proportion of section with outside rumble strip may result in 0.11 fewer crashes in a section. A 1-foot increase in the width of inside center barrier may result in 0.55 more crashes in a section. Finally, a 1-mile increase in the distance to the downstream exit ramp may result in 4.1 more crashes.

Figure 6 presents the predicted counts by random parameters negative binomial plotted against the actual observed counts. The random parameters model can decently predict the crash counts until 90 crashes but it consistently underpredicts the counts there after for higher crash counts.

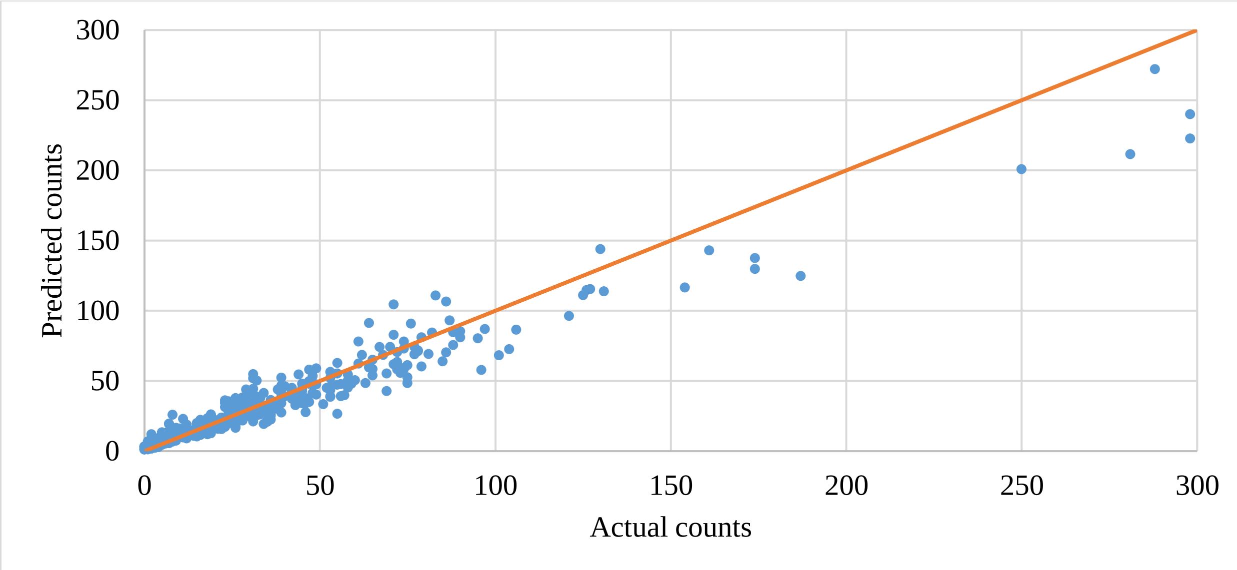

Figure 6. Observed counts versus predicted counts by random parameters negative binomial model with heterogeneity in mean for total crashes.

Estimation results of bivariate random parameters negative binomial model are presented in Table 23. Two variables, shoulder width and distance from the nearest downstream exit ramp, are found to have normally distributed random parameters for both non-injury and injury crashes. The mean of the parameter estimate for shoulder width is negative suggesting that an increase in the shoulder width reduces the frequencies of both non-injury and injury crashes. Similarly, with an increase in the distance from downstream exit ramp increases the frequencies of non-injury and injury crashes. Like the univariate model for total crashes, the parameter estimates are close to one for logarithm of section length and logarithm of number of years the section is active.

When it comes to geographical states, Hawaii is more likely to have fewer non-injury crashes when compared to Georgia and Minnesota. However, Ohio and Virginia states are found to have higher non-injury and injury crashes that those in Georgia and Minnesota states. As rumble strips are widely known to reduce the crashes, Table 23 suggests that increase in the proportion of section with rumble strips tend to reduce both the non-injury and injury crashes. Moreover, an

Table 23. Estimation results of random parameters bivariate negative binomial model for non-injury and injury crash counts on interstates.

Variables	Non-injury crashes		Injury crashes	
	Coefficient (t-statistic)	**Marginal effect**	**Coefficient (t-statistic)**	**Marginal effect**
Constant	-4.01 (-7.16)	-	-6.05 (-8.83)	-
Random parameters (normally distributed)				
Shoulder width in feet	-0.03 (-2.87)	-0.39	-0.03 (-2.59)	-0.08
Standard deviation of parameter estimate	0.05 (14.69)	-	0.04 (10.19)	-
Distance from the nearest downstream exit ramp in miles	0.01 (0.20)	0.87	0.01 (0.25)	0.08
Standard deviation of parameter estimate	0.32 (9.46)	-	0.23 (6.18)	-
Fixed parameters				
Logarithm of lane average annual daily traffic	1.38 (15.79)	0.20	1.32 (12.32)	0.03
Logarithm of section length in miles	0.98 (18.03)	3.28	0.95 (16.04)	0.56
Logarithm of number of years the section is active	1.08 (5.83)	10.57	1.34 (5.75)	2.45
Hawaii state indicator (1 if the section is in Hawaii, 0 otherwise)	-1.42 (-8.55)	-14.89	-	-
Ohio state indicator (1 if the section is in Ohio, 0 otherwise)	1.17 (10.98)	31.17	1.14 (9.87)	5.57
Virginia state indicator (1 if the section is in Virginia, 0 otherwise)	0.46 (5.59)	9.52	0.70 (8.1)	2.67
Proportion of section with outside rumble strip	-0.43 (-4.04)	-0.08	-0.47 (-4.24)	-0.016
Width of inside center barrier in feet	0.02 (1.93)	0.38	-	-
Logarithm of upstream entrance ramp average annual daily traffic	-	-	0.09 (1.95)	1.04
Overdispersion parameter	2.92 (11.93)	-	2.95 (9.86)	-
Number of observations	478			
Log-likelihood at convergence	-3094.048			
Log-likelihood with constant only	-3473.833			
Cross-equation error correlation (t-statistic)	0.816 (189.98)			

increase in the width of center barrier tend to increase the likelihood of non-injury crashes and have no significant impact on injury crashes. Finally, an increase in the average annual daily traffic of upstream entrance ramp may likely increase the number of injury crashes and have no significant impact on non-injury crashes. Marginal effects are presented in Table 23 and they are computed with same increments as those in the univariate random parameters model for total crashes in Table 22. Figure 7 and Figure 8 shows the actual observed counts versus the predicted crash counts for non-injury and injury crashes respectively. Like the univariate model for total crashes, the bivariate model is decently predicting the crashes for smaller counts and consistently under predicts for the higher crash counts. Figure 7 and Figure 8 shows the frequency distribution of differences between predicted crash counts and actual crash counts. Both the figures suggest that bivariate model is slightly over predicting the counts for both non-injury and injury crashes.

Cross-equation error correlation of 0.816 with a very high t-statistic suggests that considering bivariate models over two univariate models is more appropriate. However, two separate univariate models are estimated for non-injury and injury crashes as shown in Table 24 and Table 25 respectively. The predicted versus observed count for non-injury and injury crash counts are presented in Figure 9 and Figure 10, respectively. Both figures suggest that univariate models are decently predicting only for the non-injury and has poor performance for predicting the injury crashes. Figure 11 and Figure 12 presents the frequency distribution of predicted minus observed counts for non-injury and injury crash counts respectively. According to both figures, univariate models is over predicting both the non-injury and injury crash counts.

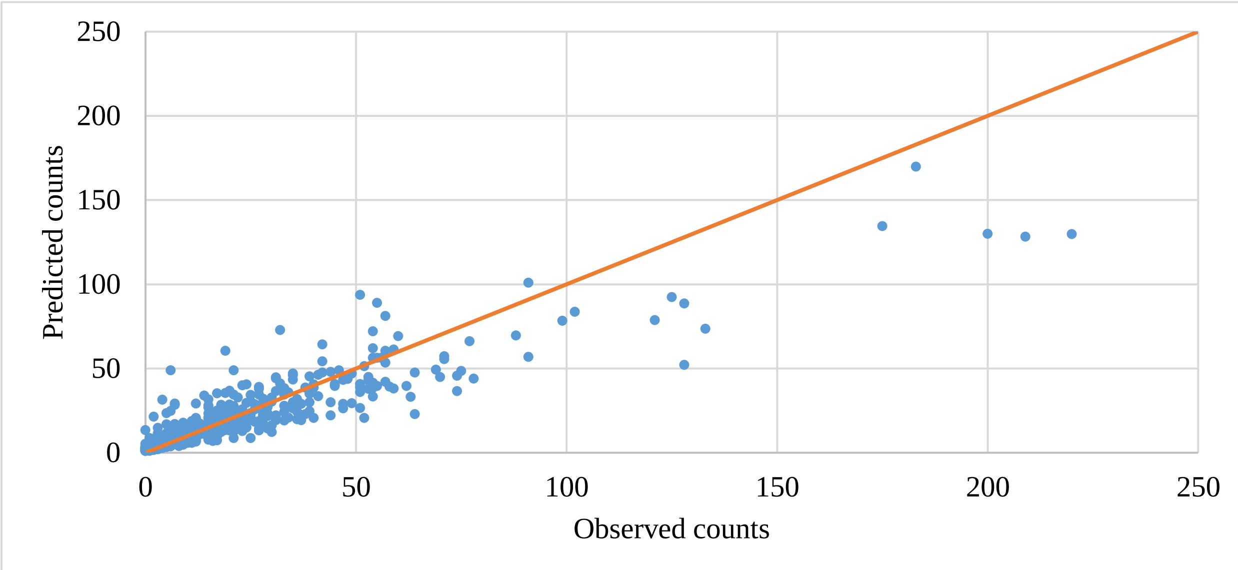

Figure 7: Observed counts versus predicted counts by random parameter bivariate negative binomial model for non-injury crashes.

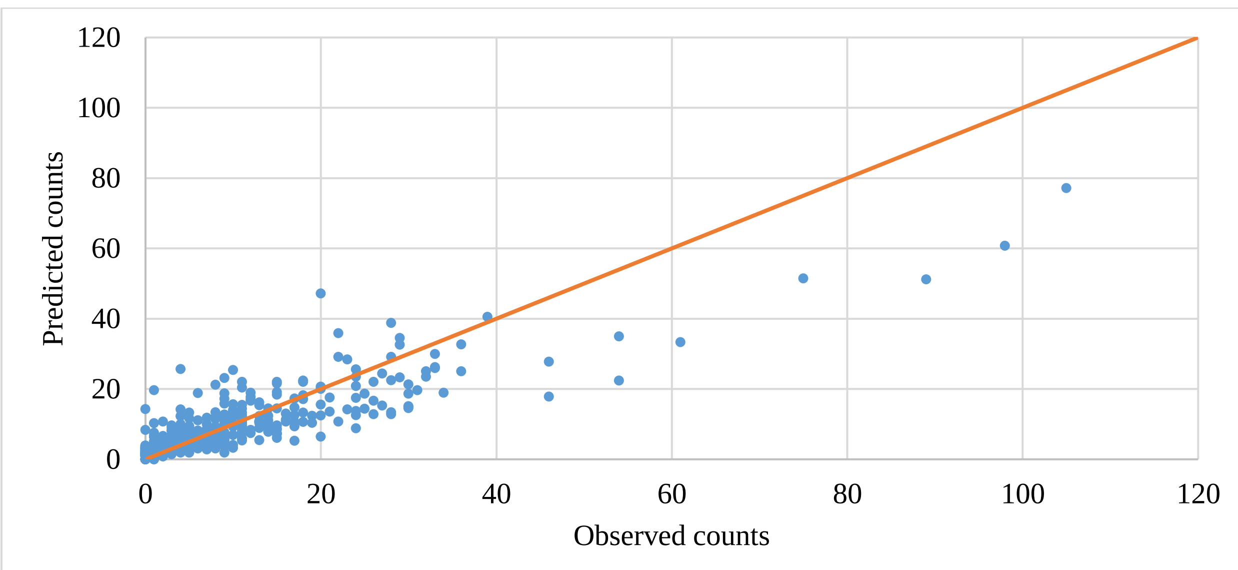

Figure 8. Observed counts versus predicted counts by random parameters bivariate negative binomial model for injury crashes.

Table 24. Estimation results of Random Parameters Negative binomial model heterogeneity in mean for non-injury crash counts on interstates.

Variables	Parameter Estimates	t-statistic	Marginal effects
Random parameter (normally distributed)			
Constant	-3.459	-8.79	
(Standard deviation of parameter distribution)	(0.611)	(25.77)	
Heterogeneity in the mean of random parameter			
Constant: Georgia state indicator (1 if the section is in Georgia, 0 otherwise)	-1.132	-8.36	
Constant: Hawaii state indicator (1 if the section is in Hawaii, 0 otherwise)	-1.303	-5.30	
Random parameter (normally distributed)			
Shoulder width in feet	-0.039	-4.57	-0.59
(Standard deviation of parameter distribution)	(0.036)	(15.98)	
Heterogeneity in the mean of random parameter			
Shoulder width in feet; Georgia state indicator (1 if the section is in Georgia, 0 otherwise)	0.042	2.68	
Shoulder width in feet; Hawaii state indicator (1 if the section is in Hawaii, 0 otherwise)	-0.089	-3.00	
Fixed parameters			
Logarithm of section length in miles	1.095	27.61	7.76
Logarithm of number of years the section is active	0.947	7.11	4.13
Logarithm of lane average annual daily traffic	1.599	25.19	1.31
Ohio state indicator (1 if the section is in Ohio, 0 otherwise)	0.697	8.83	17.69
Proportion of section with outside rumble strips	-0.369	-4.66	-0.075
Width of inside center barrier in feet	0.019	2.71	0.41
Logarithm of distance from the nearest downstream exit ramp in miles	0.091	4.44	2.49
Dispersion parameter	7.353	9.57	
Number of observations		478	
Log-likelihood at convergence		-1740.03	
Constants only log-likelihood		-1941.57	

Table 25. Estimation results of Random Parameters Negative binomial model with heterogeneity in means for injury crash counts on interstates.

Variables	Parameter Estimates	t-statistic	Marginal effects
Constant	-5.853	-8.30	
Random parameter (normally distributed)			
Shoulder width in feet (Standard deviation of parameter distribution)	-0.012 (0.025)	-0.93 (6.39)	-0.52
Heterogeneity in the mean of random parameter			
Shoulder width in feet: Georgia state indicator (1 if the section is in Georgia, 0 otherwise)	-0.054	-3.73	
Shoulder width in feet: Hawaii state indicator (1 if the section is in Hawaii, 0 otherwise)	-0.097	-4.30	
Fixed parameters			
Logarithm of section length in miles	0.973	14.40	3.99
Logarithm of number of years the section is active	1.389	5.41	3.73
Logarithm of lane average annual daily traffic	1.745	15.64	0.68
Ohio state indicator (1 if the section is in Ohio, 0 otherwise)	0.528	4.11	7.72
Proportion of section with outside rumble strip	-0.391	-3.15	-0.048
Logarithm of distance from the nearest downstream exit ramp in miles	0.107	3.25	1.71
Dispersion parameter	2.13	10.83	
Number of observations		478	
Log-likelihood at convergence		-1352.16	
Constants only log-likelihood		-1532.26	

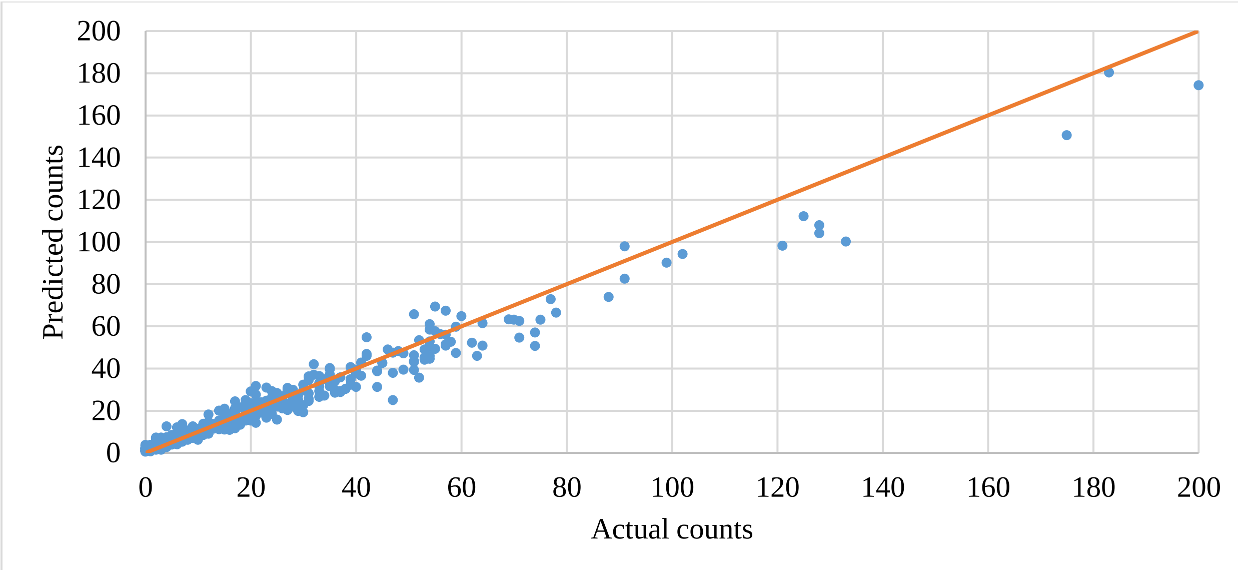

Figure 9. Observed counts versus predicted counts by random parameters negative binomial model for non-injury crashes.

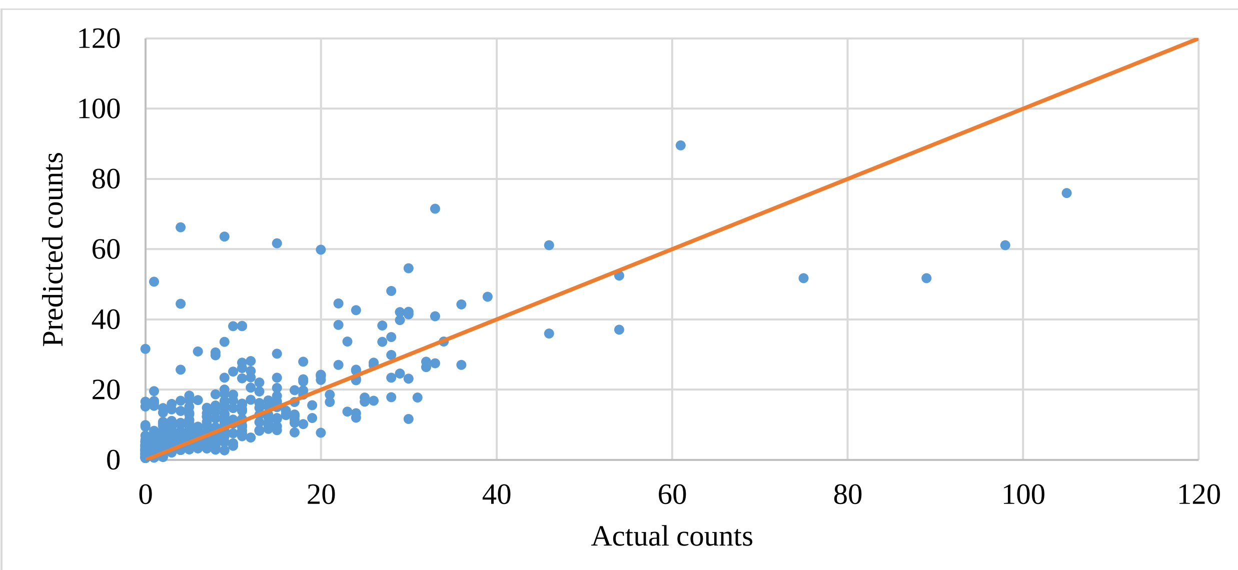

Figure 10. Observed counts versus predicted counts by random parameters negative binomial model for injury crashes.

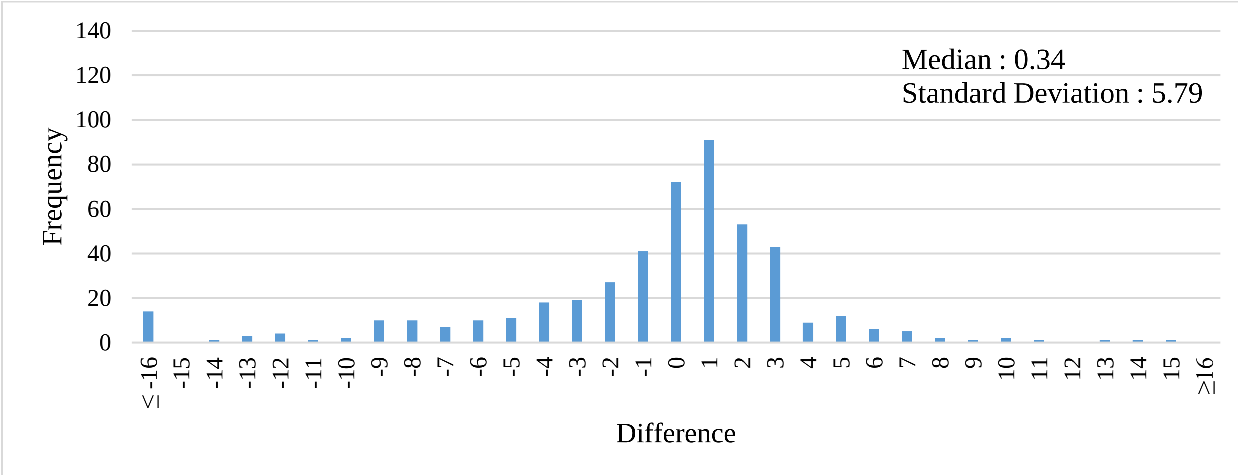

Figure 11. Frequency distribution of difference in predicted non-injury crash counts by univariate negative binomial model and observed crash counts.

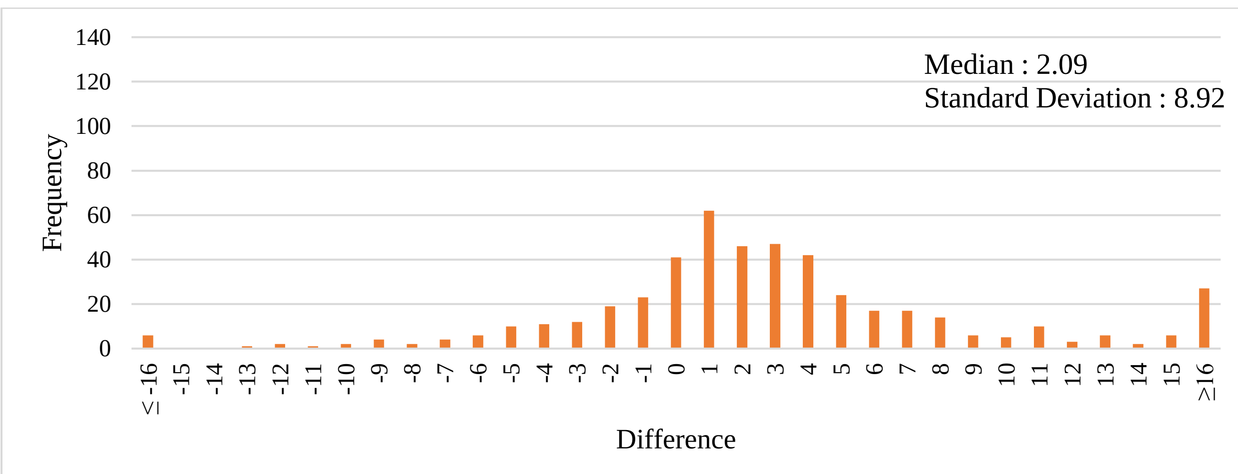

Figure 12. Frequency distribution of difference in predicted injury crash counts by univariate negative binomial model and observed crash counts.

The bivariate model performance is then compared with the univariate models using various metrics including median absolute deviation (MAD), sum of squared errors (SSE), mean squared errors (MSE), root mean squared error (RMSE), mean absolute percentage error for sections with non-zero crash counts (MAPE). Table 26 contains all the metrics calculated for bivariate and univariate models. Interestingly univariate model is performing better in all metrics only for the non-injury crashes when compared to the bivariate model. However, bivariate model is outperforming the univariate models for injury crashes. Figure 13 and Figure 14 shows the frequency distributions of crash counts prediction by bivariate model minus predictions by univariate model for non-injury and injury crashes respectively. According to Figure 13, bivariate model's count predictions for non-injury crashes are slightly higher than those by the univariate model. However, Figure 14 suggests that bivariate model's injury crash counts predictions are lower than those by the corresponding univariate model.

Table 26. Summary of goodness-of-fit measures for bivariate and univariate model predictions.

Goodness of fit test	**Non-injury crashes**		**Injury crashes**	
	Bivariate	**Univariate**	**Bivariate**	**Univariate**
Median absolute deviation (MAD)	7.36	1.84	3.71	3.33
Sum of squared errors (SSE)	80664.60	16398.69	18495.94	43345.19
Mean squared error (MSE)	168.75	34.31	38.69	90.68
Root mean squared error (RMSE)	12.99	5.86	6.22	9.52
Mean absolute percentage error (MAPE) (non-zero)	63.57	30.81	58.87	131.83

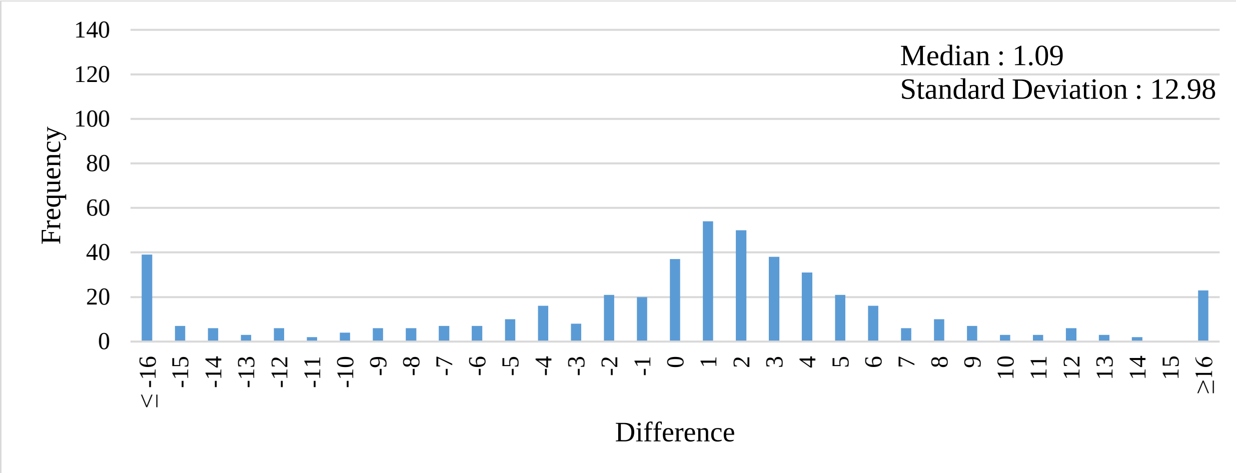

Figure 13. Frequency distribution of difference in predicted non-injury crash counts by bivariate negative binomial model and observed crash counts.

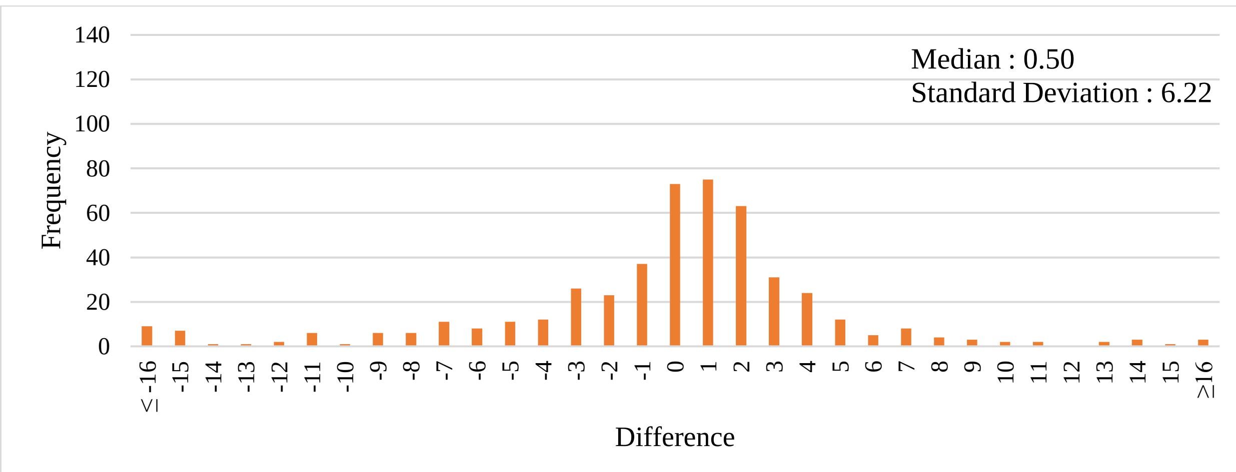

Figure 14. Frequency distribution of difference in predicted injury crash counts by bivariate negative binomial model and observed crash counts.

2.5 Bivariate/Univariate Findings Summary

Based on the performance metrics and prediction patterns, none of the estimated univariate and bivariate models were performing well for both no-injury and injury crash frequencies. Univariate models had better performance for no-injury crash frequencies and bivariate model had better performance for injury crashes. It is important to note, that the findings from this portion of the study are data specific and using a different dataset might lead to different conclusions. The data used in this study contain only 478 observations and increasing the sample size may affect the findings. The performance metrics used in this study compared the performance of models using in-sample predicted values and further studies needs to be carried out using out-of-sample predicted values. Moreover, performance metrics used in this study are based on individual level predicted values and further evaluation is needed using other performance metrics based on overall data fit such as Akaike information criterion (AIC).

Most importantly, however, both univariate and bivariate safety had statistically significant random parameters, suggesting that unobserved heterogeneity is playing an important role. Again, the performance of both of the random parameters bivariate models and the random parameters univariate models suggest that their ability to capture missing attitudinal and behavioral elements through mixing-distributions is an important practical consideration in model estimation. Because attitudinal data is nearly impossible to gather with traditional crash-data source, the importance of considering unobserved heterogeneity via mixing distributions or other methods is an empirical necessity.

Conclusions and policy implications

Highway accidents are complex events that involve a variety of human responses to external stimuli, as well as complex interactions between the vehicle, roadway features/condition, traffic-related factors, and environmental conditions. Among these complexities, attitudes toward safety play an important role and there is a vast body of evidence from fields such as psychology, neuroscience, economics, and cognitive science that suggest these attitudes change over time (Mannering, 2018). The most widely used data source (police-reported crash data) does not gather information on safety attitudes. For statistical models that are conditioned on a crash having occurred (such as the models in Part I of this report) researchers are forced to use observed data such as age and gender as proxies for safety attitudes. For models that deal solely with the frequency of crashes (such as the models in Part II of the report) crash-specific information (other than frequencies of severity) are typically not used because of the high variance of information from one crash to the next. Further complicating matters is that data such as age and gender are only gathered once a crash has occurred and the people involved in crashes are not likely to be a representative sample of the motoring population (riskier motorists will be overrepresented). It is also important to note that emerging data sources from video footage and other sources will also be missing important attitudinal information and other factors.

In the academic literature, researchers have addressed these un-collectable data as unobserved heterogeneity using mixing distributions, latent classes and other methods discussed extensively in Mannering et al. (2016). Unfortunately, heterogeneity models based on mixing distributions have not found their way into safety practice primarily because researchers have incorrectly applied them, using the means of the random parameters to forecast out-of-sample data instead of enumerating through the complete random-parameter distribution. This has led some researchers to erroneously conclude that random parameter models do not predict well. Out-of-sample prediction with both crash frequency and crash severity random parameters models must be done with full simulation through the random parameters distributions as illustrated in Hou et al. (2022). With respect to the injury-severity models, recent studies by Islam et al. (2020), Alogaili and Mannering (2020), Alnawmasi and Mannering (2022), Islam and Mannering (2022) have all used estimated random parameters injury-severity models for out-of-sample prediction by correctly simulating through the random parameter distributions, as opposed to simply using the mean of the random parameter (this approach is also demonstrated in Part I of this report). When done correctly, it can easily be shown that random parameters models predict crash frequencies and injury severities significantly better than their traditional fixed-parameters counterparts.

The findings of this report clearly show that mixing distributions (random parameters) are a statistically viable approach of capturing unobserved effects (which include motorist attitudes toward safety). Based on the findings in this report, and the findings of a growing body of recently published research, it is essential that highway safety practice incorporate unobserved heterogeneity in their safety handbooks (such as the aforementioned Highway Safety Manual). This means that the manual would have to be accompanied by a simple software package that would be capable of simulating through the random parameter distributions to get appropriate out-of-sample predictions. This would not be an onerous task and would have the potential to significantly improve safety practice and save lives.

REFERENCES

AASHTO, 2010. The Highway Safety Manual, American Association of State Highway Transportation Professionals, Washington, D.C.

Abay, K., 2013. Examining pedestrian-injury severity using alternative disaggregate models. Research in Transportation Economics, 43(1), 123-136.

Alhajyaseen, W., Iryo-Asano, M., 2017. Studying critical pedestrian behavioral changes for the safety assessment at signalized crosswalks. Safety Science, 91, 351-360.

Al-Bdairi, N., Behnood, A., Hernandez, S., 2020. Temporal stability of driver injury severities in animal-vehicle collisions: A random parameters with heterogeneity in means (and variances) approach. Analytic Methods in Accident Research 26,100120

Alnawmasi, N., Mannering, F., 2019. A statistical assessment of temporal instability in the factors determining motorcyclist injury severities. Analytic Methods in Accident Research, 22, 100090.

Alnawmasi, N., Mannering, F., 2022. The impact of higher speed limits on the frequency and severity of freeway crashes: Accounting for temporal shifts and unobserved heterogeneity. Analytic Methods in Accident Research 34, 100205.

Alogaili, A., Mannering, F., 2020. Unobserved heterogeneity and the effects of driver nationality on crash injury severities in Saudi Arabia. Accident Analysis and Prevention 144, 105618.

Anastasopoulos, P., Mannering, F., 2009. A note on modeling vehicle accident frequencies with random-parameters count models. Accident Analysis and Prevention, 41(1), 153-159.

Ariannezhad, A., Wu, Y., 2019. Effects of heavy rainfall in different light conditions on crash severity during Arizona's monsoon season. Journal of Transportation Safety and Security 11(6), 579-594.

Aziz, H., Ukkusuri, S., Hasan, S., 2013. Exploring the determinants of pedestrian–vehicle crash severity in New York City. Accident Analysis and Prevention 50, 1298-1309.

Barua, S., El-Basyouny, K., Islam, M., 2016. Multivariate random parameters collision count data models with spatial heterogeneity. Analytic Methods in Accident Research, 9, 1-15.

Batouli, G., Guo, M., Janson, B., Marshall, W., 2020. Analysis of pedestrian-vehicle crash injury severity factors in Colorado 2006–2016. Accident Analysis and Prevention 148, 105782.

Behnood, A., Mannering, F., 2016. An empirical assessment of the effects of economic recessions on pedestrian-injury crashes using mixed and latent-class models. Analytic Methods in Accident Research 12, 1-17.

Behnood, A., Mannering, F., 2019. Time-of-day variations and temporal instability of factors affecting injury severities in large-truck crashes. Analytic Methods in Accident Research 23, 1-17 100102.

Bhowmik, T., Yasmin, S., Eluru, N., 2019. Do we need multivariate modeling approaches to model crash frequency by crash types? A panel mixed approach to modeling crash frequency by crash types. Analytic Methods in Accident Research, 24, 100107.

Caliendo, C., De Guglielmo, M., Guida, M., 2016. Comparison and analysis of road tunnel traffic accident frequencies and rates using random-parameter models. Journal of Transportation Safety and Security, 8(2), 177-195.

Caliendo, C., De Guglielmo, M., Russo, I., 2019. Analysis of crash frequency in motorway tunnels based on a correlated random-parameters approach. Tunneling and Underground Space Technology, 85, 243-251.

Chen, S., Saeed, T., Labi, S., 2017. Impact of road-surface condition on rural highway safety: A multivariate random parameters negative binomial approach. Analytic Methods in Accident Research, 16, 75-89.

Chen, Z., Fan, W., 2019. Modeling pedestrian injury severity in pedestrian-vehicle crashes in rural and urban areas: mixed logit model approach. Transportation Research Record 2673(4), 1023-1034.

Coruh, E., Bilgic, A., Tortum, A., 2015. Accident analysis with aggregated data: The random parameters negative binomial panel count data model. Analytic Methods in Accident Research, 7, 37-49.

Coruh, E., Bilgic, A., Tortum, A., 2015. Accident analysis with aggregated data: The random parameters negative binomial panel count data model. Analytic Methods in Accident Research, 7, 37-49.

Dong, C., Clarke, D., Yan, X., Khattak, A., Huang, B., 2014. Multivariate random-parameters zero-inflated negative binomial regression model: An application to estimate crash frequencies at intersections. Accident Analysis and Prevention, 70, 320-329.

Dong, C., Shi, J., Huang, B., Chen, X., Ma, Z., 2017. A study of factors affecting intersection crash frequencies using random-parameter multivariate zero-inflated models. International Journal of Injury Control and Safety Promotion, 24(2), 208-221.

Eluru, N., Bhat, C., Hensher, D., 2008. A mixed generalized ordered response model for examining pedestrian and bicyclist injury severity level in traffic crashes. Accident Analysis and Prevention 40(3), 1033-1054.

FARS, 2019. Fatality Analysis Reporting System (FARS), National Highway Traffic Safety Administration (NHTSA). URL ⟨http://www-fars.nhtsa.dot.gov/Main/index.aspx⟩ (accessed 7.9.21).

Forbes, J., and Habib, M., 2015. Pedestrian injury severity levels in the Halifax regional municipality, Nova Scotia, Canada: hierarchical ordered probit modeling approach. Transportation Research Record 2519, 172-178.

Fountas, G., Anastasopoulos, P., 2018. Analysis of accident injury-severity outcomes: The zero-inflated hierarchical ordered probit model with correlated disturbances. Analytic Methods in Accident Research 20, 30-45.

Fountas, G., Fonzone, A., Gharavi, N., Rye, T., 2020. The joint effect of weather and lighting conditions on injury severities of single-vehicle accidents. Analytic Methods in Accident Research 27, 100124.

GHSA, 2020. New Projection: 2019 Pedestrian Fatalities Highest Since 1988, Governors Highway Safety Association (GHSA), URL ⟨https://www.ghsa.org/resources/news-releases/pedestrians20⟩ (accessed 7.11.21).

Green, M., 2020. Roadway human factors: from science to application. Second Edition. Lawyers and Judges Publishing Company, Tucson, AZ.

Greene, W., 2004. Interpreting estimated parameters and measuring individual heterogeneity in random parameter models. New York University Working Paper No. EC-04-08. New York, NY.

Haleem, K., Alluri, P., Gan, A., 2015. Analyzing pedestrian crash injury severity at signalized and non-signalized locations. Accident Analysis and Prevention 81, 14-23.

Hao, W., Kamga, C., Wan, D., 2016. The effect of time of day on driver's injury severity at highway-rail grade crossings in the United States. Journal of Traffic and Transportation Engineering (English edition) 3(1), 37-50.

Hou, Q., Huo, X., Leng, J., Mannering, F., 2022. A note on out-of-sample prediction, marginal effects computations, and temporal testing with random parameters crash-injury severity models. Analytic Methods in Accident Research 33, 100191.

Hou, Q., Huo, X., Tarko, A., Leng, J., 2021. Comparative analysis of alternative random parameters count data models in highway safety. Analytic Methods in Accident Research 30, 100158.

Hou, Q., Tarko, A., Meng, X., 2018. Analyzing crash frequency in freeway tunnels: a correlated random parameters approach. Accident Analysis and Prevention, 111, 94-100.

Huo, X., Leng, J., Hou, Q., Yang, H., 2020. A correlated random parameters model with heterogeneity in means to account for unobserved heterogeneity in crash frequency analysis. Transportation Research Record, 2674(7), 312-322.

Huo, X., Leng, J., Hou, Q., Zheng, L., Zhao, L., 2020. Assessing the explanatory and predictive performance of a random parameters count model with heterogeneity in means and variances. Accident Analysis and Prevention, 147, 105759.

Intini, P., Berloco, N., Fonzone, A., Fountas, G., Ranieri, V., 2020. The influence of traffic, geometric and context variables on urban crash types: A grouped random parameter multinomial logit approach. Analytic Methods in Accident Research 28 100141.

Islam, M., Alnawmasi, N., and Mannering, F., 2020. Unobserved heterogeneity and temporal instability in the analysis of work-zone crash-injury severities. Analytic Methods in Accident Research 28, 100130.

Islam, M., Mannering, F., 2022. An empirical analysis of how asleep/fatigued driving injury severities have changed over time. Journal of Transportation Safety and Security, https://doi.org/10.1080/19439962.2022.2070812.

Jägerbrand, A., Sjöbergh, J., 2016. Effects of weather conditions, light conditions, and road lighting on vehicle speed. SpringerPlus, 5(1), 1-17.

Kamla, J., Parry, T., Dawson, A., 2016. Roundabout accident prediction model: random-parameter negative binomial approach. Transportation Research Record, 2585(1), 11-19.

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

Lee, J., Mannering, F., 2002. Impact of roadside features on the frequency and severity of run-off-roadway accidents: an empirical analysis. Accident Analysis and Prevention, 34(2), 149-161.

Leone, M., Slezak, D., Golombek, D., Sigman, M., 2017. Time to decide: Diurnal variations on the speed and quality of human decisions. Cognition 158, 44-55.

Li, Y., Song, L., Fan, W., 2021. Day-of-the-week variations and temporal instability of factors influencing pedestrian injury severity in pedestrian-vehicle crashes: A random parameters logit approach with heterogeneity in means and variances. Analytic Methods in Accident Research 29, 100152.

Liu, C., Zhao, M., Li, W., Sharma, A., 2018. Multivariate random parameters zero-inflated negative binomial regression for analyzing urban midblock crashes. Analytic Methods in Accident Research, 17, 32-46.

Lord, D., Washington, S., Ivan, J., 2007. Further notes on the application of zero-inflated models in highway safety. Accident Analysis and Prevention, 39(1), 53-57.

Malyshkina, N., Mannering, F., 2010. Zero-state Markov switching count-data models: An empirical assessment. Accident Analysis and Prevention, 42(1), 122-130.

Malyshkina, N., Mannering, F., Tarko, A., 2009. Markov switching negative binomial models: an application to vehicle accident frequencies. Accident Analysis and Prevention, 41(2), 217-226.

Mannering, F., 2018. Temporal instability and the analysis of highway accident data. Analytic Methods in Accident Research 17, 1-13.

Mannering, F., Bhat, C., 2014. Analytic methods in accident research: Methodological frontier and future directions. Analytic Methods in Accident Research 1, 1-22.

Mannering, F., Bhat, C., Shankar, V., Abdel-Aty, M., 2020. Big data, traditional data and the tradeoffs between prediction and causality in highway-safety analysis. Analytic Methods in Accident Research 25, 100113.

Mannering, F., Shankar, V., Bhat, C., 2016. Unobserved heterogeneity and the statistical analysis of highway accident data. Analytic Methods in Accident Research 11, 1-16.

Mohamed, M., Saunier, N., Miranda-Moreno, L., Ukkusuri, S., 2013. A clustering regression approach: A comprehensive injury severity analysis of pedestrian–vehicle crashes in New York, US and Montreal, Canada. Safety Science 54, 27-37.

Mohammadi, M., Samaranayake, V., Bham, G., 2014. Crash frequency modeling using negative binomial models: An application of generalized estimating equation to longitudinal data. Analytic Methods in Accident Research, 2, 52-69.

Mokhtarimousavi, S., 2019. A time of day analysis of pedestrian-involved crashes in California: Investigation of injury severity, a logistic regression and machine learning approach using HSIS data. ITE Journal 89(10), 25-33.

Mokhtarimousavi, S., Anderson, J., Azizinamini, A., Hadi, M., 2020. Factors affecting injury severity in vehicle-pedestrian crashes: a day-of-week analysis using random parameter ordered response models and artificial neural networks. International Journal of Transportation Science and Technology 9(2), 100-115.

Naznin, F., Currie, G., Logan, D., Sarvi, M., 2016. Application of a random effects negative binomial model to examine tram-involved crash frequency on route sections in Melbourne, Australia. Accident Analysis and Prevention, 92, 15-21.

Pantangi, S., Ahmed, S., Fountas, G., Majka, K., Anastasopoulos, P., 2021. Do high visibility crosswalks improve pedestrian safety? A correlated grouped random parameters approach using naturalistic driving study data. Analytic Methods in Accident Research 30, 100155.

Peltzman, S., 1975. The effects of automobile safety regulation, Journal of Political Economy 83(4), 677–725.

Plainis, S., Murray, I., 2002. Reaction times as an index of visual conspicuity when driving at night. Ophthalmic and Physiological Optics 22(5), 409-415.

Pour-Rouholamin, M., Zhou, H., 2016. Investigating the risk factors associated with pedestrian injury severity in Illinois. Journal of Safety Research 57, 9-17.

Qiu, B., Fan, W., 2021. Mixed logit models for examining pedestrian injury severities at intersection and non-intersection locations. Forthcoming in Journal of Transportation Safety and Security, https://doi.org/10.1080/19439962.2021.1923101.

Rusli, R., Haque, M., Afghari, A., King, M., 2018. Applying a random parameters Negative Binomial Lindley model to examine multi-vehicle crashes along rural mountainous highways in Malaysia. Accident Analysis and Prevention, 119, 80-90.

Rusli, R., Haque, M., King, M., Voon, W.,2017. Single-vehicle crashes along rural mountainous highways in Malaysia: an application of random parameters negative binomial model. Accident Analysis and Prevention, 102, 153-164.

Saeed, T., Hall, T., Baroud, H., Volovski, M., 2019. Analyzing road crash frequencies with uncorrelated and correlated random-parameters count models: An empirical assessment of multilane highways. Analytic Methods in Accident Research, 23, 100101.

Savolainen, P., Mannering, F., Lord, D., Quddus, M., 2011. The statistical analysis of crash-injury severities: A review and assessment of methodological alternatives. Accident Analysis and Prevention 43(5), 1666-1676.

Shankar, V., Milton, J., Mannering, F., 1997. Modeling accident frequencies as zero-altered probability processes: an empirical inquiry. Accident Analysis and Prevention, 29(6), 829-837.

Shaon, M., Qin, X., Shirazi, M., Lord, D., Geedipally, S., 2018. Developing a random parameters negative binomial-Lindley model to analyze highly over-dispersed crash count data. Analytic Methods in Accident Research, 18, 33-44.

Se, C., Champahom, T., Jomnonkwao, S., Karoonsoontawong, A., Ratanavaraha, V., 2021. Temporal stability of factors influencing driver-injury severities in single-vehicle crashes: A correlated random parameters with heterogeneity in means and variances approach. Analytic Methods in Accident Research 32,100179.

Song, L., Li, Y., Fan, W., Liy, P., 2021. Modeling pedestrian-injury severities in pedestrian-vehicle crashes considering spatiotemporal patterns: Insights from different hierarchical Bayesian random-effects models. Traffic Injury Prevention 22(7), 524-529.

Song, L., Li, Y., Fan, W., Wu, P., 2020. Modeling pedestrian-injury severities in pedestrian-vehicle crashes considering spatiotemporal patterns: Insights from different hierarchical Bayesian random-effects models. Analytic Methods in Accident Research 28, 100137.

Sullivan, J., Flannagan, M., 2011. Differences in geometry of pedestrian crashes in daylight and darkness. Journal of Safety Research 42(1), 33-37.

Sullivan, J., 2001. Characteristics of pedestrian risk in darkness. University of Michigan, Ann Arbor, Transportation Research Institute. Ann Arbor, MI.

Tang, F., Fu, X., Cai, M., Lu, Y., Zhong, S., Lu, C., 2020. Applying a correlated random parameter negative binomial Lindley model to examine crash frequency along highway tunnels in China. IEEE Access, 8, 213473-213488.

Tyrrell, R., Brooks, J., Wood, J., Carberry, T., 2004. Nighttime conspicuity from the pedestrian's perspective. In 83rd Annual Meeting of the Transportation Research Board, Washington, DC.

Ukkusuri, S., Hasan, S., Aziz, H., 2011. Random parameter model used to explain effects of built-environment characteristics on pedestrian crash frequency. Transportation research record, 2237(1), 98-106.

Venkataraman, N., Shankar, V., Ulfarsson, G., Deptuch, D., 2014. A heterogeneity-in-means count model for evaluating the effects of interchange type on heterogeneous influences of interstate geometrics on crash frequencies. Analytic Methods in Accident Research, 2, 12-20.

Venkataraman, N., Ulfarsson, G., Shankar, V., 2013. Random parameter models of interstate crash frequencies by severity, number of vehicles involved, collision and location type. Accident Analysis and Prevention, 59, 309-318.

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. Transportation Research Record, 2236(1), 41-48.

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. Transportation Research Record, 2236(1), 41-48.

Wang, C., Chen, F., Cheng, J., Bo, W., Zhang, P., Hou, M., Xiao, F., 2020. Random-Parameter Multivariate Negative Binomial Regression for Modeling Impacts of Contributing Factors on the Crash Frequency by Crash Types. Discrete Dynamics in Nature and Society, 2020.

Washington, S., Karlaftis, M., Mannering, F., Anastasopoulos, P., 2020. Statistical and Econometric Methods for Transportation Data Analysis. CRC Press.

Winkelmann, R., 2008. Econometric analysis of count data. Springer Science and Business Media.

Winston, C., Maheshri, V., Mannering, F., 2006. An exploration of the offset hypothesis using disaggregate data: The case of airbags and antilock brakes. Journal of Risk and Uncertainty, 32(2), 83-99.

WHO, 2018. Global status report on road safety 2018, World Health Organization (WHO), URL ⟨http://apps.who.int/iris/bitstream/handle/10665/277370/WHO-NMH-NVI-18.20-eng.pdf?ua=1⟩ (accessed 7.9.21).

Xin, C., Guo, R., Wang, Z., Lu, Q., Lin, P., 2017. The effects of neighborhood characteristics and the built environment on pedestrian injury severity: a random parameters generalized ordered probability model with heterogeneity in means and variances. Analytic Methods in Accident Research, 16, 117-132.

Xu, P., Zhou, H., Wong, S.C., 2021. On random-parameter count models for out-of-sample crash prediction: Accounting for the variances of random-parameter distributions. Accident Analysis and Prevention 159, 106237.

Yan, X., He, J., Zhang, C., Liu, Z., Wang, C., Qiao, B., 2021. Temporal analysis of crash severities involving male and female drivers: A random parameters approach with heterogeneity in means and variances. Analytic Methods in Accident Research 30,100161.

Yasmin, S., Eluru, N., Ukkusuri, S., 2014. Alternative ordered response frameworks for examining pedestrian injury severity in New York City. Journal of Transportation Safety and Security, 6(4), 275-300.

Yu, M., Zheng, C., Ma, C., 2020. Analysis of injury severity of rear-end crashes in work zones: A random parameters approach with heterogeneity in means and variances. Analytic Methods in Accident Research 27,100126.

Zamani, A., Behnood, A., Davoodi, S.R., 2021. Temporal stability of pedestrian injury severity in pedestrian-vehicle crashes: New insights from random parameter logit model with heterogeneity in means and variances, Analytic Methods in Accident Research, 32, 100184.

Zhai, X., Huang, H., Sze, N., Song, Z., Hon, K., 2019. Diagnostic analysis of the effects of weather condition on pedestrian crash severity. Accident Analysis and Prevention, 122, 318-324.