SAFETY ANALYSIS OF HIGHWAY MERGING AND DIVERGING AREAS USING ADVANCED ECONOMETRIC METHODS

PENGLIN SONG

PhD

The Hong Kong Polytechnic University

2024

The Hong Kong Polytechnic University

Department of Civil and Environmental Engineering

SAFETY ANALYSIS OF HIGHWAY MERGING AND DIVERGING AREAS USING ADVANCED ECONOMETRIC METHODS

PENGLIN SONG

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

May 2024

CERTIFICATE OF ORIGINALITY

I hereby declare that this thesis is my own work and that, to the best of my knowledge and belief, it reproduces no material previously published or written, nor material that has been accepted for the award of any other degree or diploma, except where due acknowledgement has been made in the text.

_________________________ (Signed)

Penglin SONG (Name of student)

Abstract

Considering the substantial social costs of road crashes, it is important to identify the road safety hazards at crash-prone locations, such as transition segments where vehicles merge onto or exit the main highway. Therefore, effective road safety measures can be developed to mitigate the crash risk. Conventional highway safety studies rely on statistical models based on historical crash data. However, they may have limitations because of the rare and random nature of crash occurrences. Additionally, some crucial information like dynamic traffic characteristics is often not available in the crash dataset. Furthermore, there are shortfalls in existing statistical models. Therefore, this study aims to develop advanced statistical and econometric methods for the safety analysis of highway merging and diverging areas, which have relatively high crash risk among all road entities.

First, effects of possible influencing factors on the crash injury severity are different between single-vehicle and multi-vehicle crashes. Therefore, separate crash injury severity models should be established for single- and multi-vehicle crashes. Additionally, transferability assessment of the crash injury severity models across different time periods should be considered. To this end, the random parameter multinomial logit regression model with heterogeneity in means and variances is adopted to measure the association between crash injury severity and possible influencing factors at the highway ramp areas (including acceleration and deceleration lanes), with which the effect of unobserved heterogeneity is accounted for. Furthermore, partially constrained and temporal unconstrained modeling approaches are adopted to account for temporally shifting parameters. Results indicate that there are considerable differences in the effects of aberrant driver behavior, vehicle class, lighting condition and crash location on injury severity between single- and multi-vehicle crashes. Last but not least, out-of-sample prediction performance for the

models of single- and multi-vehicle crashes are assessed.

Second, a key assumption for multiple regression model is the exogeneity of the independent variables. However, some influencing factors that affect crash injury severity may correlate with other unobserved factors. As such, there could exist an endogenous effect of crash type on crash injury severity. Hence, a random parameter recursive bivariate probit model is proposed to model the crash type (hit-object and rollover) and crash injury severity at ramps simultaneously and to account for any endogenous effect of crash type. The results indicate that the proposed model is capable of capturing the correlation between unobserved factors and crash type. Endogeneity effect of crash type is incorporated in the crash severity model using a correlation parameter. Furthermore, other exogenous variables including road environment, crash location, and driver characteristics that affect the crash type and crash injury severity at the highway ramp areas were also identified.

Third, information on dynamic traffic characteristics is often not available in historical crash dataset. To resolve this problem, vehicle trajectories can be captured using videos and image processing techniques in real-time safety analysis. In this study, a modified conflict risk indicator is developed to assess the safety of diverging, merging, and weaving movements of traffic, with which vehicles' dimensions (width and length), and longitudinal and angular movements are considered. A correlated grouped random parameter logit model with heterogeneity in the means is established to capture the unobserved heterogeneity, with additional flexibility, at road user level for the association between conflict risk and influencing factors. Prevalence and severity of both rear-end and sideswipe traffic conflicts are examined. Results indicate that toll collection type, vehicle location, average longitudinal speed, angular speed, acceleration, and vehicle class all affect the risk of traffic conflicts.

Overall, contribution of this study is twofold. First, advanced statistical and econometric methods are developed to account for several prevalent issues in safety analysis. Second, alternate safety indicators are adopted to proxy the crash risk at the highway merging and diverging areas. Findings should shed light on effective traffic management and control measures like variable speed limits and message signs that can mitigate the crash risk at the hazardous locations.

List of Publications

Publications arising from the thesis

Song, P., Sze, N.N.*, Guo, J., Zhu, D. (2024) Temporal transferability assessment of injury severity models for single-vehicle and multi-vehicle crashes at highway ramp areas accounting for unobserved heterogeneity. Under review.

Song, P., Sze, N.N.*, Chen, S., Labi, S. (2024) Correcting for endogeneity of crash type in crash injury severity at highway ramp areas. *Accident Analysis and Prevention* 208, 107785.

Song, P., Sze, N.N.*, Zheng, O., Abdel-Aty, M. (2022) Addressing unobserved heterogeneity at road user level for the analysis of conflict risk at tunnel toll plaza: A correlated grouped random parameters logit approach with heterogeneity in means. *Analytic Methods in Accident Research* 36, 100243.

Conference proceedings/presentations

Song, P., Sze, N.N., Chen, S., Labi, S. (2024) Correcting for endogeneity between crash injury severity and crash type at freeway ramp areas using a hierarchical Bayesian bivariate ordered approach. Paper presentation at the Transportation Research Board 103rd Annual Meeting, 7-11 January, Washington, D.C., United States.

Song, P., Sze, N.N. (2023) Correcting for endogeneity between crash injury severity and crash type at freeway ramp areas using a hierarchical Bayesian bivariate ordered approach. Paper presented at the 27th International Conference of Hong Kong Society for Transportation Studies, 11-12 December, Hong Kong.

Song, P., Sze, N.N., Guo, J. (2023) Investigating differences between injury severity of single-vehicle and multi-vehicle crashes at freeway ramp areas. Paper presentation at the Transportation Research Board 102nd Annual Meeting. 8-12 January, Washington, D.C., United States.

Song, P., Sze, N.N. (2022) A random parameters approach with heterogeneity in means and variances for crash-injury severity at highway ramp areas. Paper presentation at the 26th International Conference of Hong Kong Society for Transportation Studies, 12-13 December, Hong Kong.

Song, P., Sze, N.N. (2021) Conflict-based safety analysis at tunnel toll plaza. Paper presentation at the 25th International Conference of Hong Kong Society for Transportation Studies, 9-10 December, Hong Kong.

Acknowledgements

I am deeply grateful to those who have supported and guided me throughout my research journey.

First and foremost, I would like to express my heartfelt gratitude to my chief supervisor, Dr. Nang-Ngai (Tony) Sze. His invaluable guidance, unwavering support, and insightful feedback have been instrumental in shaping my research. His patience and encouragement have kept me motivated through the research journey. I know that I may not have been the best student, but he has undoubtedly been the best supervisor for me, always understanding and supportive.

I would like to extend my sincere thanks to my co-supervisor, Prof. Zhen Leng. His encouragement during my undergraduate years was pivotal in my decision to pursue research. I am grateful for his inspiration to chase my research ambitions and for guiding me with his expertise.

I am profoundly thankful to Prof. Samuel Labi and Dr. Sikai Chen for their support and guidance during my visit to Purdue University and University of Wisconsin–Madison. Their insights and mentorship greatly enhanced my research experience and broadened my academic perspective. I am also grateful to the group members at both institutions for their camaraderie and assistance.

I am grateful to all members of our research group, colleagues, and friends at PolyU, for their companionship and support. I am thankful for the stimulating conversations and the unwavering support.

To my family, words cannot express my gratitude for their endless love and encouragement. To my parents, Mr. Jian Song and Ms. Junying Dong, I am deeply thankful for their sacrifices and for always believing in me. Their constant support has been my anchor throughout this journey. Lastly, I would like to thank Ms. Huimin Hu for her love, patience, and encouragement. Her support has been my source of strength and motivation.

Table of Contents

Abstract i

List of Publications iv

Acknowledgements vi

Table of Contents viii

List of Figures xi

List of Tables xii

Chapter 1 Introduction 1

1.1 Research background 1

1.2 Motivation and problem statement 2

1.3 Research objectives 4

1.4 Thesis organization 5

Chapter 2 Literature Review 7

2.1 Safety analysis at merging and diverging areas 7

2.2 Surrogate safety measures 10

2.3 Analytical methodology and frontiers 12

2.3.1 Discrete outcome model 12

2.3.2 Unobserved heterogeneity 13

2.3.3 Transferability and temporal instability 15

2.3.4 Endogeneity 16

2.4 Concluding remarks 18

Chapter 3 Temporal Transferability of Crash Injury Severity Models 20

3.1 Introduction 20

3.2 Data 22

3.3 Method 31

3.3.1 Modeling approach 31

3.3.2 Transferability assessment 32

3.4 Results and discussion 34
3.4.1 Driver characteristics 35
3.4.2 Vehicle attributes..36
3.4.3 Environmental conditions 37
3.4.4 Roadway design 38
3.4.5 Crash circumstances..39
3.4.6 Heterogeneity in means and variances..39
3.5 Transferability assessment 45
3.5.1 Temporal stability 45
3.5.2 Out-of-sample prediction 46
3.6 Concluding remarks 48
Chapter 4 Correcting for Endogeneity of Crash Type in Crash Injury Severity at Highway Ramp Areas 50
4.1 Introduction..50
4.2 Data 51
4.3 Method 55
4.4 Results and discussion 58
4.4.1 Effect of endogeneity 58
4.4.2 Influencing factors of crash type..59
4.4.3 Influencing factors of injury severity..60
4.5 Concluding remarks 69
Chapter 5 Addressing Unobserved Heterogeneity at Road User Level for the Analysis of Conflict Risk at Toll Plaza 71
5.1 Introduction..71
5.2 Method 72
5.2.1 Traffic conflict 72
5.2.2 Model formulation 75
5.3 Data 78

5.4 Results and discussion 83
5.4.1 Toll payment types 84
5.4.2 Vehicle speed 85
5.4.3 Vehicle acceleration rate 85
5.4.4 Vehicle class 86
5.4.5 Spatial location 87
5.4.6 Unobserved heterogeneity 87
5.5 Concluding remarks 92
Chapter 6 Conclusions and Recommendations 94
6.1 Conclusions 94
6.2 Findings and contributions 97
6.3 Limitations and recommendations 98
6.3.1 Limitations 98
6.3.2 Recommendations for future research 99
References 101
Appendix 122

List of Figures

Figure 1.1 Police-reported motor vehicle crashes at ramps in US from 2018 to 2021 2

Figure 2.1 Safety pyramid of traffic events 11

Figure 3.1 Police-reported motor vehicle crashes in US from 2018 to 2021 21

Figure 3.2 Illustration of typical ramp areas 23

Figure 4.1 Distribution of random parameters 68

Figure 5.1 Illustration of interaction between two conflicting vehicles 73

Figure 5.2 Illustration of possible conflict scenarios 74

Figure 5.3 Layout of study site 80

Figure 5.4 Vehicle trajectories for different toll payment types 81

Figure 5.5 Vehicle trajectories for different vehicle classes 82

List of Tables

Table 2-1 Factors that affect the crash injury severity at ramp areas............................8

Table 3-1 Descriptive statistics of the data for single-vehicle crashes 24

Table 3-2 Descriptive statistics of the data for multi-vehicle crashes 27

Table 3-3 Results of partially constrained parameter estimation for single-vehicle crashes..41

Table 3-4 Results of partially constrained parameter estimation for multi-vehicle crashes..43

Table 3-5 Results of likelihood ratio tests for temporal stability of single-vehicle crash 45

Table 3-6 Results of likelihood ratio tests for temporal stability of multi-vehicle crash 46

Table 3-7 Difference in the average predicted probabilities for single vehicle crash 47

Table 3-8 Difference in the average predicted probabilities for multi-vehicle crash...47

Table 4-1 Descriptive statistics of the data 53

Table 4-2 Cross tabulation of injury severity and crash type......................................54

Table 4-3 Results of parameter estimation...64

Table 4-4 Marginal effects for injury 68

Table 5-1 Distributions of the sample by toll payment type and vehicle class 80

Table 5-2 Descriptive statistics of variables considered 83

Table 5-3 Model performance metrics between uncorrelated and correlated models 84

Table 5-4 Results of parameter estimation of correlated model with heterogeneity in the means for rear-end conflicts...89

Table 5-5 Results of parameter estimation of correlated model with heterogeneity in

the means for sideswipe conflicts 90
Table 5-6 Cholesky matrix of random parameters for rear-end conflict (t-statistic in parentheses)..91
Table 5-7 Correlation coefficient matrix of random parameters for rear-end conflict 91
Table 5-8 Cholesky matrix of random parameter for sideswipe conflict (t-statistic in parentheses)..91
Table 5-9 Correlation coefficient matrix of random parameters for sideswipe conflict 91
Table 5-10 Marginal effects on the probabilities of rear-end conflicts and sideswipe conflicts..91

Chapter 1 Introduction

1.1 Research background

Highway safety is of paramount importance due to its direct impact on transportation systems, public health, economic costs, and overall societal well-being. Research has shown that traffic accidents on highways result in a significant number of injuries and fatalities, leading to immense human suffering and loss. The economic burden of these accidents is substantial, encompassing costs related to medical expenses, property damage, legal fees, and lost productivity. Considering the substantial social costs of road crashes, it is important to identify the road safety hazards at crash-prone locations.

Highway merging and diverging areas, such as ramps and toll plazas, serve as transition segments where vehicles merge onto or exit the main highway, leading to potential interactions and conflicts that can increase the probability of accidents. Often, vehicles on the road at one level need to ascend or descend to the crossing road at a different level, and this is done via highway ramps. Ramp areas tend to be accident-prone because ramp traffic consists of a mix of vehicles based on their intentions: continuing to go straight, changing lanes so the vehicle can be positioned to be ready to exit or enter the mainline via the ramp, and so on. These different motives lead to a complex tapestry of vehicle movements (weaving, merging, diverging) and consequently a wide variation in speeds. At such locations, there exists challenges and difficulties in the driving tasks, navigation, and maneuvers including acceleration and deceleration, diverging and merging, lane changing, and gap acceptance. Further, traffic conflicts between vehicles of different streams are prevalent. For these reasons, crash risk at highway ramps tends to be high at high-speed road corridors in several countries (***Wang et al., 2009***; ***Geedipally and Lord,***

2010; ***Li et al., 2012***). For example, as shown in **Figure 1.1**, in the United Staes, in recent four years, the annual average crashes at highway ramps is 271,000 resulting in over 68,000 injury crashes and 800 fatal crashes per year (***NHTSA, 2020***, ***2021***, ***2022***, ***2023***). The risk of overall and severe crashes at highway ramp area is higher than that at other road entities. Therefore, from a safety management perspective, it can be beneficial to identify the factors that contribute to the high crash risk at ramp areas, measure the strength of their influence.

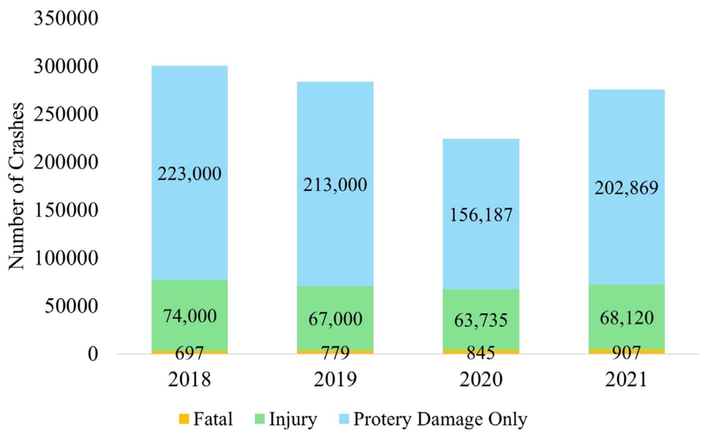


Figure 1.1 Police-reported motor vehicle crashes at ramps in US from 2018 to 2021

Toll plaza is another type of merging and diverging area for toll collection of controlled-access roads like expressways, bridges, and tunnels. Traffic and safety characteristics of toll plazas are different from that of other road entities because of the differences in geometric design, traffic management and control, and more importantly, weaving, diverging, and merging movements of traffic approaching the toll booths, especially vehicles slow down or stop to pay tolls when multiple toll collection methods (i.e., manual, and electronic) are available (***Wong et al., 2006***), posing unique safety challenges that require careful management to prevent collisions.

1.2 Motivation and problem statement

The safety analysis of highway merging and diverging areas is essential. The specific research motivation and problem statement of this study are as follows.

First, previous studies indicate that differences in the characteristics of crash occurrence and severity between single- and multi-vehicle crashes are considerable (***Martensen and Dupont, 2013***; ***Rezapour et al., 2019***). The crash mechanism and effects of influencing factors between single- and multi-vehicle crashes are different (***Mannering and Bhat, 2014***; ***Intini et al., 2020***). Indeed, effectiveness of road management strategies and measures in reducing the crash and injury risk could be diminished if the crash mechanisms of single-vehicle and multi-vehicle crashes are not differentiated (***Intini et al., 2020***), especially for highway merging and diverging areas. Therefore, it is necessary to evaluate the differences in the association measure of crash severity between single-vehicle and multi-vehicle crashes.

Second, transferability allows for the application of models to be effectively used for new data from other temporal and/or spatial units, enhancing the generalizability and applicability of the findings. Nevertheless, the transferability remains a crucial issue since the effects of the same explanatory variable are likely to be temporally and/or spatially unstable (***Mannering and Bhat, 2014***; ***Mannering, 2018***). Therefore, the transferability of crash injury severity models is worth studying.

Third, there are unobserved and unmeasurable factors that may affect the association between observed variables of interest and outcome variables in safety analysis. Variations in traffic operations, vehicle maneuvers and driver behavior, which are unobservable, in the diverging, merging, and weaving process at the ramp areas should be considerable. Ignoring such effects may result in erroneous inferences (***Savolainen et al., 2011***; ***Mannering and Bhat, 2014***). Therefore, the effect of unobserved heterogeneity should be considered.

Moreover, crash type was typically incorporated into the crash severity model as an independent input variable in previous studies. However, as is the case in any crash data, there could exist unobserved factors, including driver perception and behaviors, which may affect both the crash type and crash severity simultaneously. In other words, the endogenous effect of crash type on crash injury severity may lead to biased parameter estimates. Therefore, it is important to accommodate the endogenous effect in the analysis of road safety.

Last but not least, previous studies are based on historical crash records to address heterogeneity issues, not many studies have considered the heterogeneity and correlation in real-time conflict risk estimation at the road user level. Also, the association between crash occurrence and possible risk factors can be moderated by collision type. Therefore, considering the conflict type and heterogeneity at the road user level could provide new insights for safety analysis.

1.3 Research objectives

This research aims to assess safety at highway merging and diverging areas using advanced econometric methods, with specific research objectives as follows:

- Transferability of crash injury severity models for single-vehicle and multi-vehicle crashes

1) To investigate the differences in the association measure of injury severity at ramp areas between single-vehicle and multi-vehicle crashes, for which effects of unobserved heterogeneity are accounted for.
2) To address the issues of transferability over time by considering temporally shifting parameters for the analysis of crash injury severity at ramps.

- Correcting for crash type endogeneity in crash injury severity at highway ramps

1) To investigate the crash type and injury severity simultaneously.
2) To examine the endogenous effect of crash type and indirect effects of exogenous factors on injury severity through crash types.

- Addressing unobserved heterogeneity at road user level for the analysis of conflict risk at toll plaza

1) To assess the safety risk at a tunnel toll plaza diverging area using a modified traffic conflict indicator, taking into account vehicle length and width, angular and longitudinal movements, and conflict type (i.e., rear-end and sideswipe).
2) To examine the association between conflict risk at tunnel toll plaza and possible factors, including vehicle class, speed and acceleration of vehicle, toll collection type, and spatial characteristics, for which effects of unobserved heterogeneity and correlation among random parameters at the road user level are accounted for.

1.4 Thesis organization

This thesis is organized into six comprehensive chapters, each dedicated to a specific aspect of the study.

Chapter 2 reviews the literature on highway safety analysis including safety analysis based on crash data, surrogate safety measures and traffic conflict technique for safety analysis and advanced analytic methods and critical methodological issues relating to highway safety analysis.

Chapter 3 investigates differences in the association measure of injury severity at ramp areas between single-vehicle and multi-vehicle crashes. The issues of unobserved heterogeneity and temporal instability for the analysis of crash injury severity at ramp areas are addressed. In this study, random parameters multinomial logit regression approach, with heterogeneity in means and variances, is adopted to measure the association between possible influencing factors and crash injury severity at ramps.

Chapter 4 develops a random parameters recursive bivariate probit model to investigate the crash type and injury severity simultaneously, using single-vehicle crash data at ramps. In the proposed simultaneous model, crash type is regarded as the treatment variable. Additionally, the effect of unobserved heterogeneity is also considered using the random parameters model with heterogeneity in the means.

Chapter 5 proposes a modified traffic conflict indicator, taking into account vehicle length and width, angular and longitudinal movements, and conflict type (i.e., rear-end and sideswipe), is proposed to assess the safety risk at a tunnel toll plaza, based on high-resolution vehicle trajectory data obtained from drone video. Then, the correlated grouped random parameter multinomial logit approach with heterogeneity in the means of the random parameters is adopted to measure the association between conflict risk at tunnel toll plaza and possible factors, including vehicle class, speed and acceleration of vehicle, toll collection type, and spatial characteristics, for which effects of unobserved heterogeneity and correlation among random parameters at the road user level are accounted for.

Chapter 6 concludes the study with a summary of findings, contributions, limitations and future research directions.

Chapter 2 Literature Review

This chapter reviews the literature on highway safety analysis from several aspects. **Section 2.1** reviews the literature on the highway safety analysis based on crash data, with a focus on crash injury severity analysis at merging and diverging areas. **Section 2.2** discusses the surrogate safety measures and traffic conflict technique for safety analysis. **Section 2.3** reviews the literature with respect to analytic methods and critical methodological issues relating to highway safety analysis.

2.1 Safety analysis at merging and diverging areas

The safety analysis using highway accident data relied on the data from police crash reports. These reports are used to establish the frequency of crashes at specific locations and the associated injury-severities of vehicle occupants and others involved in these crashes (***Mannering and Bhat, 2014***). Studies have assessed the safety of merging and diverging areas (e.g., ramps, toll plazas) based on historical crash data. For example, toll plaza layout, horizontal curves, toll collection method, and traffic signs and road markings are found associated with the crash risk at toll plazas (***Wong et al., 2006***; ***Sze et al., 2008***; ***Abuzwidah et al., 2014***; ***Abuzwidah and Abdel-Aty, 2015***, ***2018***). Factors that affect the crash injury severity at ramp areas can be classified into categories including driver, vehicle, environmental, roadway design, and temporal characteristics. **Table 2-1** summarizes the main findings of the factors that affect the crash injury severity at ramp areas.

Additionally, some studies also indicated that the prediction performances of separated crash severity models for single- and multi-vehicle crashes are superior to that of combined models (***Geedipally and Lord, 2010***; ***Mannering, 2018***). However, studies that consider the difference in the effects of influencing factors on the crash injury severity between these two crash types are limited (***Savolainen and Mannering, 2007***; ***Yu and Abdel-Aty, 2013***; ***Wu et al., 2014***; ***Gong et al., 2022***). Moreover, it is rare that the issues of unobserved heterogeneity and temporal instability are addressed for the crash severity analysis at ramp areas.

Table 2-1 Factors that affect the crash injury severity at ramp areas

Factor	**Findings**
Driver characteristics	
Gender	Likelihood of severe injury of **female driver** is higher (***Mergia et al., 2013***).
Age	Effect of **driver age** on injury severity varies across models for merging and diverging areas (***Mergia et al., 2013***).
Driving under the influences of alcohol and drug	Likelihoods of severe injury and fatal injury are higher when **driving under the influence of alcohol and drug** (***Wang et al., 2009***; ***Li et al., 2012***; ***Mergia et al., 2013; Zhang et al., 2018***).
Aberrant driver behavior	Likelihood of severe injury of **speeding offense** is higher (***Mergia et al., 2013***). Likelihood of injury of **improper lane change** is higher at merging and merging overlap influence areas (***Yang et al., 2019***).
Vehicle attributes	
Vehicle class	Likelihood of no injury of **heavy vehicle** is higher (***Wang et al., 2009***). Likelihood of severe injury of **semi-truck** is higher at merging areas, and likelihood of non-incapacitating injury of **semi-truck** is lower at diverging areas (***Mergia et al., 2013***). Likelihood of injury of **truck** is lower at merge and weaving overlap influence areas (***Yang et al., 2019***).
Environmental conditions	
Road surface condition	Likelihood of severe injury in **wet road surface condition** is lower (***Wang et al., 2009***; ***Li et al., 2012***; ***Zhang et al., 2018***), but the likelihood of invisible injury is higher (***Zhang et al., 2018***). The likelihoods of possible injury, non-incapacitating injury and incapacitating injury are higher in **poor road condition,** but the likelihood of fatal injury is lower (***Mergia et al., 2013***).
Lighting condition	Likelihood of injury is lower under **daylight condition** (***Wang et al., 2009***; ***Li et al., 2012***; ***Zhang et al., 2018***). Likelihood of non-incapacitating injury at diverging areas is higher under **poor lighting condition** (***Mergia et al., 2013***). Likelihood of injury at diverging and diverging overlap influence areas is higher under **poor lighting condition** (***Yang et al., 2019***).
Weather	Likelihood of severe injury is lower under **clear weather condition** (***Wang et al., 2009***; ***Li et al., 2012***; ***Zhang et al., 2018***). Effect of **adverse weather conditions** varies across estimation models (***Mergia et al., 2013***). Likelihood of injury at diverging areas is higher under **adverse weather condition** (***Yang et***

Factor	Findings
	al., 2019).
Land use	Likelihoods of no injury and severe injury are higher at **commercial area** (***Wang et al., 2009***; ***Li et al., 2012***; ***Zhang et al., 2018***). Likelihood of severe injury is higher at **rural area** (***Geedipally et al., 2014***).
Roadway design	
Speed limit	**Posted speed limit of mainline** is positively associated with the likelihood of no injury (***Wang et al., 2009***).
Traffic volume	**Traffic volume of mainline** is positively associated with the likelihood of severe injury (***Wang et al., 2009***). **Truck volume of mainline** and **traffic volume of exit ramps** are positively associated with the likelihood of severe injury of truck-related crash (***Wang et al., 2011***). **Traffic volume of diverging area** is negatively associated with the likelihood of fatal injury and **traffic volume of merging area** is positively associated with the likelihood of incapacitating injury (***Mergia et al., 2013***).
Horizontal and vertical alignments	Likelihood of severe injury is higher when the road alignment is **horizontal curve** and **vertical grade** (***Wang et al., 2009***; ***Wang et al., 2011***).
Width of central median	Likelihood of severe injury decreases when the **central median** is wide (***Wang et al., 2011***).
Width of shoulder	**Wide right shoulder (left-hand drive)** is negatively associated with the likelihood of no injury (***Wang et al., 2009***), but positively associated with the likelihood of injury (***Zhang et al., 2018***). The likelihood of severe injury of truck-related crash at **wide shoulder** area (left-hand drive) is lower (***Wang et al., 2011***). The impact of **wide right shoulder of freeway mainline** (left-hand drive) on injury severity varies among different methods (***Li et al., 2012***).
Number of lanes	**More mainline lanes** and **more ramp lanes** are positively associated with the likelihood of fatal injury at merging areas; **more ramp lanes** are positively associated with the likelihood of non-incapacitating injury at diverging areas, but negatively associated with the likelihood of severe injury (***Mergia et al., 2013***; ***Geedipally et al., 2014***). **More mainline lanes** are positively associated with the likelihood of injury (***Wang et al., 2009***; ***Wang et al., 2011***; ***Li et al., 2012***; ***Zhang et al., 2018***).
Ramp configuration	Likelihood of severe injury of **long deceleration lane** is lower (***Wang et al., 2009***; ***Wang et al., 2011***). Likelihood of severe injury of **long exit ramp** is higher (***Wang et al., 2009***). Effect of **long exit ramp** on injury severity varies among different methods ***Li et al., 2012***; ***Zhang et al., 2018***).
Road barrier	Likelihoods of severe injury and non-incapacitated injury decrease with the **presence of barrier** (***Geedipally et al., 2014***).
Crash circumstances	
Collision type	Likelihood of injury of **sideswipe crash** is lower (***Wang et al., 2009***; ***Li et al.,***

Factor	Findings
	2012; ***Zhang et al., 2018***). Likelihood of severe injury of **rear end crash** is lower (***Li et al., 2012***; ***Zhang et al., 2018***), but likelihood of invisible injury of that is higher (***Zhang et al., 2018***). Likelihoods of no injury and possible injury of **rear end crash** or **same direction sideswipe crash** are higher, and likelihood of fatal injury of **angle crash** is higher (***Mergia et al., 2013***).
Crash location	Likelihood of severe injury at **exit ramp area** is higher (***Geedipally et al., 2014***).
Temporal characteristics	
Time of day	Likelihood of severe injury at **off-peak period** is higher (***Wang et al., 2009***).

2.2 Surrogate safety measures

Road safety analysis based on historical crash data is often subject to the problems like under-reporting, misclassification, and imbalanced crash data (***Tsui et al., 2009***; ***Lord and Mannering, 2010***; ***Savolainen et al., 2011***; ***Mannering and Bhat, 2014***). To this end, it is possible to estimate the safety risk based on real-time traffic data collected using video observational survey and driving simulator approaches (***Sayed et al., 2013***; ***Yun et al., 2017***; ***Chen et al., 2019b***; ***Saad et al., 2019***; ***Arun et al., 2021b***; ***Wang et al., 2022***; ***Chen et al., 2024***). Surrogate safety measures like time and distance headway, mean and deviation of speed, acceleration rate, and traffic conflicts can be applied to assess the safety level of road entities. To estimate the risk of traffic conflicts, indicators like time-to-collision, post-encroachment time, and deceleration rate to avoid the crash are used (***Tarko, 2018***). According to the "safety pyramid", as shown in **Figure 2.1**, traffic incidents can be classified into three categories: (i) normal interactions; (ii) traffic conflicts; and (iii) crashes (***Hydén, 1987***). As a crash is the extreme form of traffic conflict, modeling the latter (which requires shorter observation period to accrue enough sample) may provide a reliable foundation for better understanding of crash mechanisms (***El-Basyouny and Sayed, 2013***; ***Sayed et al., 2013***; ***Zheng et al., 2021***). Despite that, more work is required for accurate prediction of crash severity based on traffic conflict analysis (***Paul and Ghosh, 2021***).

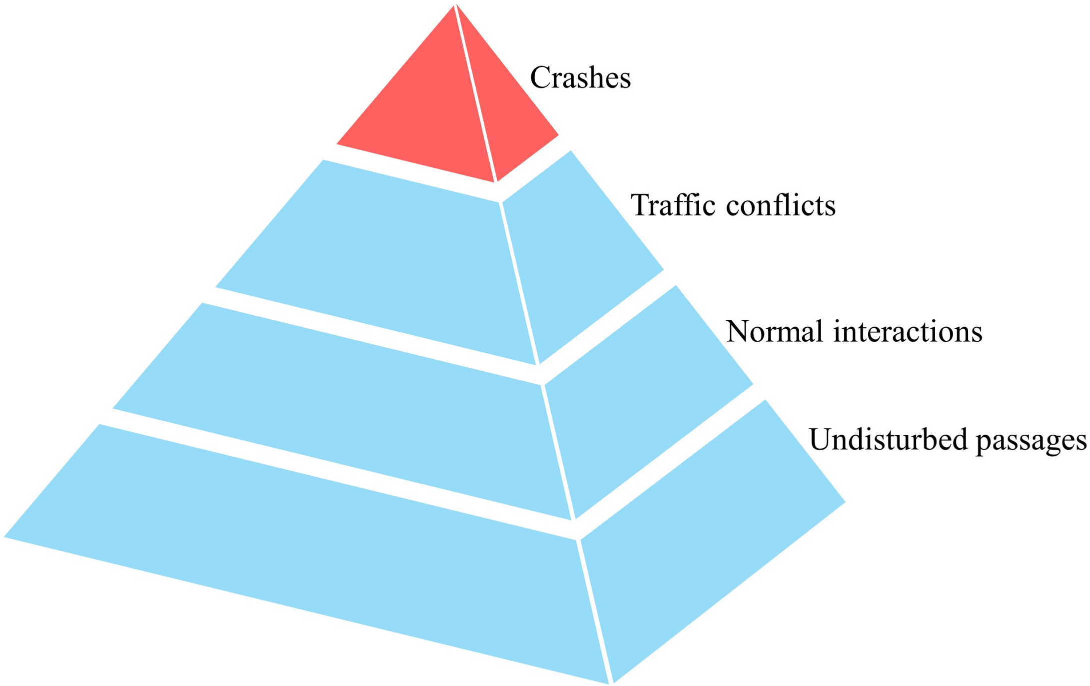


Figure 2.1 Safety pyramid of traffic events

In most of the previous studies, the risk of traffic conflicts has been estimated based on vehicle length, position of centroid, and longitudinal movement of conflicting vehicles only (***Sacchi and Sayed, 2016***; ***Wang et al., 2019***; ***Xu et al., 2021***; ***Yang et al., 2021***). It may result in underestimation of traffic conflict risk and bias in parameter estimates when vehicle width, and motions in two dimensions of conflicting vehicles are not considered. This is particularly true for the interactions between vehicles at intersections, and diverging, merging, and weaving areas. To this end, it is necessary to consider vehicle width and length (***Arun et al., 2021c***; ***Arun et al., 2021a***), two-dimensional (i.e., longitudinal and angular) movements (***Ward et al., 2015***; ***Tarko, 2021***), and point of contact (***Laureshyn et al., 2010***; ***Jiménez et al., 2013***) of conflicting vehicles in traffic conflict analysis.

In recent years, it is increasingly popular to collect traffic data using unmanned aerial vehicles (also known as drones) for traffic and safety analysis (***Stipancic et al., 2016***;

Wang et al., 2019). Based on the aerial footage captured by the drone, it is possible to extract high-resolution vehicle trajectory data using automated image recognition and processing technique (***Krajewski et al., 2018***; ***Li et al., 2020***). Although several studies have assessed the safety risk of toll plazas based on drone data, they do not explicitly account for the effects of vehicle dimensions and conflict type on risk estimation (***Xing et al., 2019***; ***Xing et al., 2020a***; ***Xing et al., 2020b***). For example, it is possible to distinguish among various conflict types, i.e., head-on, sideswipe, and angled conflicts using pixel-based image classification technique based on high resolution vehicle trajectory data (***Wu et al., 2020***).

2.3 Analytical methodology and frontiers

2.3.1 Discrete outcome model

For the analytic methods, discrete outcome models[1] are commonly used to model the crash injury severity since the crash injury severity levels are generally classified into discrete categories in the crash dataset (***Savolainen et al., 2011***; ***Mannering and Bhat, 2014***). The injury severity levels are usually either a binary discrete outcome (e.g., injury and non-injury) or a multiple discrete outcome (e.g., no injury, possible injury, evident injury, disabling injury, fatal injury). For example, a binary logit model is applied to examine the association between influencing factors and a dichotomous dependent variable (killed or severe injury against slight injury) to evaluate the pedestrian injury risk (***Sze and Wong, 2007***). Dependent variables with multiple discrete outcomes can be considered ordered or unordered. To model the likelihood of more than two crash injury outcomes, a variety of multinomial models that do not

[1] It should be noted that when dealing with the decision maker's choices among discrete alternatives, such as a traveler's choice of mode, discrete outcome models are often referred to as discrete choice models. In the analysis of road safety, the more general term "discrete outcome model" is used for any discrete outcomes, such as crash injury severity.

account for the ordering of injury outcomes have been widely used, such as multinomial logit model (***Rifaat et al., 2011***; ***Ye and Lord, 2014***; ***Song et al., 2024b***). Modeling approaches that consider the ordinal nature of injury severity levels, such as ordered probit/logit model, have also been applied (***Abdel-Aty, 2003***; ***Lee and Abdel-Aty, 2005***; ***Mergia et al., 2013***). Moreover, the severity of traffic conflicts is also typically classified into discrete categories such as binary categories or multiple categories. For example, a previous study divided traffic conflicts measured by TTC into two categories with a threshold of 4 seconds to evaluate the crash potential (***Xing et al., 2019***). Therefore, discrete outcome models have been applied to model the association between severity of conflicts and possible factors like road geometry, environmental condition, traffic flow, and driver characteristics (***Uzondu et al., 2018***; ***Xing et al., 2019***; ***Xing et al., 2020a***)

2.3.2 Unobserved heterogeneity

In discrete outcome models, fixed parameters restrict the effects of explanatory variables to be the same across all observations. However, there are unobserved and unmeasurable factors that may affect the association between observed variables of interest and outcome variables (e.g., injury severity, conflict severity). For example, there are significant differences in (unobserved) safety perception, attitude, and travel habit among the drivers who are of the same age. Hence, driver age may not be able to fully account for the effect of unobserved individual heterogeneity on driver performance and travel behavior. Thus, parameter estimates would be biased when unobserved heterogeneity is not considered. Ignoring such effects may result in erroneous inferences (***Savolainen et al., 2011***; ***Mannering and Bhat, 2014***). To this end, various modeling approaches including random parameters, latent-class, and Markov switching approaches can be adopted to account for the effect of unobserved heterogeneity (***Train, 2009***; ***Anastasopoulos and Mannering, 2011***; ***Morgan and Mannering, 2011***; ***Ye and Lord, 2014***; ***Hensher et al., 2015***; ***Yasmin et al., 2015***).

The random parameter (effect) approach has been the most widely applied to model the severity of crash injury and traffic conflicts For example, effects of unobserved heterogeneity of passenger characteristics are considered in driver injury severity analysis ***(Behnood and Mannering, 2017)***. Also, gender, bike type, signal type, and bike volume are found to be associated with unobserved heterogeneity in bicyclists' red-light running behavior (***Guo et al., 2018***). In addition, to better track unobserved heterogeneity, it is necessary to capture heterogeneity in the standard deviations of parameter by allowing the mean of parameter to be a function of explanatory variables using heterogeneity in means and variances approach (***Hensher et al., 2015***; ***Mannering et al., 2016***). This further defines the dispersion of parameter values across individual observations, providing additional flexibility for capturing potential unobserved heterogeneity (***Seraneeprakarn et al., 2017***). For example, a random parameters logit model with heterogeneity in means was developed to consider the effect of passengers on driver injury severities (***Behnood and Mannering, 2017***). Based on this concept, several studies have considered heterogeneity in the means and variances of the random parameters to capture unobserved heterogeneity with additional flexibility (***Alnawmasi and Mannering, 2019***; ***Islam et al., 2020***; ***Yu et al., 2020***; ***Alogaili and Mannering, 2022***; ***Song et al., 2022***). Furthermore, more extensions based on random parameters models are considered. For example, correlated random parameters model is adopted to account for correlation among random parameters (***Fountas et al., 2018b***; ***Hou et al., 2020***), and grouped random parameters model is applied to indicate the unobserved effects due to repeated observations of the same entity (***Fountas et al., 2018a***). Nevertheless, model extensions include correlated random parameters model with heterogeneity in the means (***Ahmed et al., 2021***; ***Fountas et al., 2021***; ***Se et al., 2021***; ***Pantangi et al., 2022***), correlated grouped random parameters model (***Meng et al., 2021***), grouped random parameters with heterogeneity in the means (***Ahmed et al., 2020***), and correlated grouped random parameters model with heterogeneity in the means

(***Pantangi et al., 2020***). However, the aforementioned studies were based on historical crash record, not many studies have considered the heterogeneity in the means of the random parameters and correlation among random parameters in real-time conflict risk estimation at the micro-level (***Li et al., 2021***; ***Zhang et al., 2021***).

2.3.3 Transferability and temporal instability

In road safety analysis, transferability refers to the extent to which estimated parameters, findings, or methodologies developed in one context (e.g., a specific time period or region) can be applied or generalized to another context while maintaining relevance and accuracy (***Xu et al., 2014***; ***Essa et al., 2019***; ***Washington et al., 2020***; ***Arun et al., 2022***). Nevertheless, the transferability of crash injury severity model across different spatial and temporal units remains a crucial issue since the effects of the same explanatory variable are likely to be temporally and/or spatially unstable (***Mannering and Bhat, 2014***; ***Mannering, 2018***). For example, previous study has found that the effect of factors that determine injury severity in large truck crashes varies significantly at different times of the day and in different years (***Behnood and Mannering, 2019***). Recently, numerous studies have demonstrated that the contributing factors related to injury severities are subject to changes over time (***Behnood and Mannering, 2019***; ***Islam et al., 2020***; ***Islam and Mannering, 2020***; ***Meng et al., 2021***; ***Yan et al., 2021***; ***Zamani et al., 2021***; ***Alnawmasi and Mannering, 2022***, ***2023***). A series of likelihood ratio tests have been used to test the models' overall spatial and temporal transferability. Furthermore, it is necessary to examine the temporal changes in the influences of specific explanatory variables on the outcomes. To this end, temporal unconstrained and constrained parameters can be introduced into the crash injury severity model. Then, the partially temporal constrained parameters model, where some of the parameters are constrained and others are unconstrained, is considered to estimate the shifts in the effect of a specific variable over time (***Alnawmasi and Mannering, 2023***; ***Dzinyela et al., 2024***).

Therefore, bias in parameter estimation could be eliminated, and understanding on the transferability of influencing factors over time that affect the crash severity should be improved.

2.3.4 Endogeneity

Endogeneity, which is characterized by significant correlation between an explanatory variable and the error term of a regression model, can be attributed to the influence of omitted variables, measurement errors, simultaneity, and self-selection (***Mannering and Bhat, 2014***; ***Guevara, 2015***; ***Mannering et al., 2020***). Endogeneity could lead to bias and inconsistency of parameter estimation, and ultimately, faulty inferences and false conclusions. For example, considering the context of seat belt use, a driver's use of seat belt is typically considered as an exogenous variable in conventional crash severity studies. However, a driver that does not use a seat belt tends to be taking a risk intrinsically. It is likely that aggressive driving behavior is involved. Hence, the probability of more severe injury would increase. To this end, the influence of seat belt non-use on the crash severity would be overestimated if the endogenous effect were not considered (***Eluru and Bhat, 2007***). In other words, driver safety perception is an unobserved factor that could affect the likelihood of seat belt use and crash injury severity. This, possibly, contributes to the endogenous effect of a driver's non-use of a seat belt.

Several methodological approaches have been established to eliminate the bias of parameter estimation attributed to endogeneity (***Train, 2009***; ***Guevara, 2015***). For example, the control-function method with instrumental variables has been a direct way to account for endogeneity (***Guevara and Ben-Akiva, 2012***; ***Guevara and Hess, 2019***). An instrumental variable for each endogenous variable is required in the model. Such an instrumental variable is highly correlated with the endogenous variable while independent from the error term of the model. Several studies have used instrumental variables to explore the relationship among multiple sources of risk (***Afghari et al., 2018***), driver sleepiness and headway (***Afghari et al., 2022***), speed enforcement and safety risk (***Yasmin et al., 2022***), and motorcyclist behavior and injury severity (***Yu et al., 2023***). One of the major considerations for the application of the control-function method is to identify appropriate instrumental variables. This is important particularly in the context of safety analysis (***Mannering and Bhat, 2014***; ***Guevara, 2015***; ***Mannering et al., 2020***). However, identifying suitable instrumental variables for omitted attributes can be challenging in some circumstances. For example, driver safety perception and vehicle performance may be highly correlated with crash injury severity but are difficult to measure from the historical crash data. Hence, the control-function method with instrumental variable may not be appropriate for the crash severity analysis (***Chang et al., 2022***).

A simultaneous equation model (also known as simultaneous equation system or maximum likelihood approach, see ***Train (2009)***), is a statistical model with two or more equations where an endogenous variable in one equation can be estimated by a function of exogenous variables in other equations in the equation system[2] (***Greene, 2018***; ***Washington et al., 2020***). Simultaneous equation models capture the endogeneity by the cross-equation correlations and account for the indirect effects using a recursive structure. In past safety analysis, simultaneous equation models have been used in several studies to explore the endogeneity effect of possible factors in models of crash severity (***Eluru and Bhat, 2007***; ***Rana et al., 2010***; ***Li et al., 2018***; ***Chang et al., 2022***), crash frequency (***Bhat et al., 2014***; ***Heydari et al., 2020***; ***Heydari and Forrest, 2024***), driving behavior (***Sarwar et al., 2017***), risk compensation (***Oviedo-Trespalacios et al., 2020***), seat belt use (***Afghari et al., 2021***), and road user

[2] If the dependent variable in one equation does not serve as an explanatory variable in other equations, it is called bivariate (or multivariate) dependent variable models (***Washington et al., 2020***).

interactions in shared space (***Kazemzadeh and Afghari, 2024***). For example, a simultaneous equation model was adopted to account for the endogeneity effect of seat belt use on crash severity (***Eluru and Bhat, 2007***; ***Abay et al., 2013***). Also, a two-equation system – recursive bivariate probit model – was adopted to explore inter-connected choices among passenger characteristics and crash circumstances (***Lee and Abdel-Aty, 2008***). In addition, a copula-based approach was proposed to address the endogeneity problem in injury severity models for two-vehicle crashes (***Rana et al., 2010***). Furthermore, in crash studies, it is crucial to consider the effect of unobserved heterogeneity on parameter estimation (***Li et al., 2021***; ***Song et al., 2022***)(***Li et al., 2021***; ***Song et al., 2022***). Several studies have extended the simultaneous equation models for identifying any random effects of exogenous variables (***Eluru and Bhat, 2007***; ***Abay et al., 2013***; ***Chang et al., 2022***).

2.4 Concluding remarks

This chapter summarizes the literature review on safety analysis. The strengths and limitations of crash injury severity analysis, surrogate safety measures, and analytical frontiers are examined. There are several research gaps identified from previous research, which are summarized as follows.

Even that the crash mechanism and effects of influencing factors between single- and multi-vehicle crashes are different, previous studies have rarely differentiated between them (***Mannering and Bhat, 2014***; ***Intini et al., 2020***), especially for highway merging and diverging areas. Therefore, it is necessary to evaluate the differences in the association measure of crash severity between single-vehicle and multi-vehicle crashes.

Variations in traffic operations, vehicle maneuvers and driver behavior, which are unobservable, in the diverging, merging, and weaving process at the ramp areas should be considerable. The effect of unobserved heterogeneity should be considered. Nevertheless, the transferability remains a crucial issue since the effects of the same explanatory variable are likely to be temporally and/or spatially unstable (***Mannering and Bhat, 2014***; ***Mannering, 2018***).

There could exist unobserved factors, including driver perception and behaviors, which may affect both the crash type and crash severity simultaneously. In other words, the effect of crash type on crash injury severity, could be endogenous. It is rare that previous studies had considered the indirect effects of exogenous factors on injury severity.

Last but not least, the association between crash occurrence and possible risk factors can be moderated by collision type. The heterogeneity and correlation in real-time conflict risk estimation at the road user level are rarely considered in previous studies.

Chapter 3 Temporal Transferability of Crash Injury Severity Models

3.1 Introduction

As shown in **Figure 3.1**, during the four-year period from 2018 to 2021, 70% of road crashes in the United States involved more than one vehicle, and 30% involved one vehicle respectively. For instance, number of single-vehicle crashes slightly reduced from 1,934,000 in 2018 to 1,874,000 in 2019. In contrast, the number of multi-vehicle crashes increased from 4,801,000 to 4,881,000 in the same period. Nevertheless, during the COVID-19 pandemic, there were remarkable reductions in both single-vehicle (1,710,635) and multi-vehicle (3,540,202) crashes in 2020 (***NHTSA, 2020***, ***2021***, ***2022***, ***2023***). Studies indicate that differences in the characteristics of crash occurrence and severity between single- and multi-vehicle crashes are considerable (***Martensen and Dupont, 2013***; ***Rezapour et al., 2019***). Even that the crash mechanism and effects of influencing factors between single- and multi-vehicle crashes are different, it is rare that previous studies had differentiated between them (***Mannering and Bhat, 2014***; ***Intini et al., 2020***). Existing studies examined the association between influencing factors and crash injury severity of overall crashes (without differentiating between single-vehicle and multi-vehicle crashes) at ramps. Indeed, effectiveness of road management strategies and measures in reducing the crash and injury risk could be diminished if the crash mechanisms of single-vehicle and multi-vehicle crashes are not differentiated (***Intini et al., 2020***), especially for ramp areas. Therefore, it is necessary to evaluate the differences in the association measure of crash severity at ramps between single-vehicle and multi-vehicle crashes.

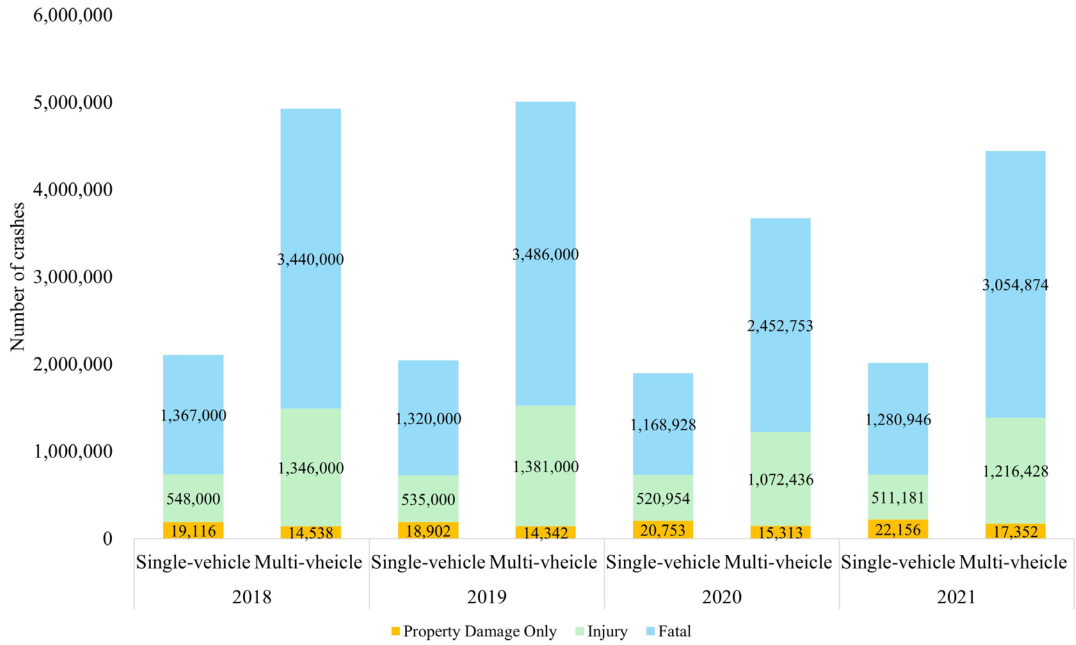


Figure 3.1 Police-reported motor vehicle crashes in US from 2018 to 2021

On the other hand, two modeling issues including unobserved heterogeneity and transferability for crash severity analysis should be addressed (***Mannering et al., 2016***; ***Behnood and Mannering, 2017***; ***Seraneeprakarn et al., 2017***; ***Xin et al., 2017***; ***Mannering, 2018***; ***Alnawmasi and Mannering, 2019***; ***Islam et al., 2020***; ***Yu et al., 2020***; ***Alogaili and Mannering, 2022***; ***Alnawmasi and Mannering, 2023***). For example, relationship between built environment and pedestrian injury severity is examined using the random parameters ordered probit model, with which the random effects of older pedestrians were revealed (***Xin et al., 2017***). In addition, it is necessary to consider the variations by time (i.e., time of the day, and over the years) of influencing factors that affect the crash severity for the optimal policy strategies and recommendations that can reduce the injury risk (***Behnood and Mannering, 2015***; ***Alogaili and Mannering, 2022***). Temporal instabilities of the association between possible influencing factors and crash injury severity at work zone were investigated (***Islam et al., 2020***; ***Yu et al., 2020***). Just, it is rare that the issues of unobserved heterogeneity and temporal instability are addressed for the analysis of crash injury severity at ramp areas.

To this end, this study aims to investigate the differences in the association measure of injury severity at ramp areas between single-vehicle and multi-vehicle crashes and address the issues of transferability over time for the analysis of crash injury severity at ramps. In this study, random parameters multinomial logit regression approach, with heterogeneity in means and variances, is adopted to measure the association between possible influencing factors and crash injury severity at ramps, based on the crash data from the State of North Carolina in 2016-2018, with which temporally shifting parameters are considered using partially constrained and unconstrained temporal models.

The remainder of this chapter is structured as follows. Description of the data used, and analysis method are given in **Section 3.2** and **Section 3.3**, respectively. **Section 3.4** presents the results of parameter estimation. **Section 3.5** discusses the transferability assessment of the models. Finally, concluding remarks are given in **Section 3.6**.

3.2 Data

Ramp areas consist of ramp proper and speed change lanes (i.e., acceleration lanes for on-ramp and deceleration lanes for off-ramp) in the United States (***AASHTO, 2018***). In this study, crash data at the ramp areas (**Figure 3.2**) in North Carolina State in the period 2016 to 2018, obtained from the Highway Safety Information System (HSIS), is used. According to the injury definition adopted by North Carolina, crash injury severity is classified into five categories, namely fatal injury, suspected serious injury, suspected minor injury, possible injury, and no injury. To avoid the problem of imbalanced crash data, fatal injury and suspected serious injury are combined into one class as "severe injury", and suspected minor injury and possible injury are combined into one class as "minor injury" respectively (***Islam et al., 2020***; ***Alogaili and***

Mannering, 2022). Hence, three discrete crash injury severity outcomes including no injury, minor injury, and severe injury would be estimated. In addition, injury severity analysis is stratified into two for single-vehicle and multi-vehicle crashes. For a crash that involves more than one personal injury, crash severity would be determined based on the victim who suffers from the most serious injury. For each crash, information on driver characteristics, vehicle attributes, environmental conditions, roadway design, and crash circumstance is available. **Table 3-1** and **Table 3-2** summarize the descriptive statistics of the sample. Overall, there are 3,170 (19.0% of all crashes, including 2294 no injuries, 813 minor injuries, and 63 severe injuries) single-vehicle and 13,541 (81.0% of all crashes, including 10613 no injuries, 2815 minor injuries, 113 severe injuries) multi-vehicle crashes at ramp areas from 2016 to 2018. Result of a chi-square test indicates that the null hypothesis that the injury severities among single-vehicle crashes and multi-vehicle crashes are the same can be rejected at the 1% level of significance (critical chi-squared value of 72.35 with 2 degree of freedom).

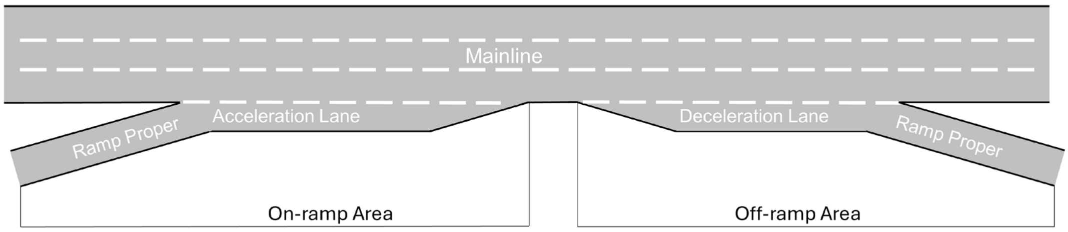


Figure 3.2 Illustration of typical ramp areas

Table 3-1 Descriptive statistics of the data for single-vehicle crashes

Variable	Attribute	2016		2017		2018	
		Count	Percentage	Count	Percentage	Count	Percentage
Injury severity	No injury	703	70.80%	765	73.91%	826	72.33%
	Minor injury	272	27.39%	245	23.67%	296	25.92%
	Severe injury	18	1.81%	25	2.42%	20	1.75%
Gender	Male	619	62.34%	680	65.70%	707	61.91%
	Female	374	37.66%	355	34.30%	435	38.09%
Age	Below 25	364	36.66%	354	34.20%	367	32.14%
	25-39	326	32.83%	360	34.78%	415	36.34%
	40-59	221	22.26%	228	22.03%	273	23.91%
	Above 59	82	8.26%	93	8.99%	87	7.62%
Alcohol or Drugs	Not under the influence of alcohol or drugs	913	91.94%	952	91.98%	1050	91.94%
	Driving under the influence of alcohol or drugs	80	8.06%	83	8.02%	92	8.06%
Aberrant driving behavior	Speeding violation	388	39.07%	359	34.69%	401	35.11%
	Oversteer	91	9.16%	117	11.30%	107	9.37%
	Inattentiveness	81	8.16%	62	5.99%	99	8.67%
	Aggressive driving	53	5.34%	63	6.09%	61	5.34%
	Other aberrant driving behavior	380	38.27%	434	41.93%	474	41.51%
Safety belt	Not used	54	5.44%	73	7.05%	75	6.57%
	Used	939	94.56%	962	92.95%	1067	93.43%
Maneuver	Changing lanes	159	16.01%	145	14.01%	194	16.99%
	Going straight	710	71.50%	743	71.79%	797	69.79%

Variable	Attribute	2016		2017		2018	
		Count	Percentage	Count	Percentage	Count	Percentage
	Making turn	81	8.16%	97	9.37%	107	9.37%
	Other maneuver action	43	4.33%	50	4.83%	44	3.85%
Vehicle type	Car	620	62.44%	635	61.35%	698	61.12%
	Sport utility vehicle	159	16.01%	180	17.39%	209	18.30%
	Pickup	106	10.67%	115	11.11%	133	11.65%
	Van	41	4.13%	37	3.57%	24	2.10%
	Truck	57	5.74%	59	5.70%	64	5.60%
	Other vehicle types	10	1.01%	9	0.87%	14	1.23%
Road surface condition	Dry	564	56.80%	617	59.61%	637	55.78%
	Wet	359	36.15%	383	37.00%	427	37.39%
	Other road surface conditions	70	7.05%	35	3.38%	78	6.83%
Lighting condition	Daylight	538	54.18%	583	56.33%	640	56.04%
	Dusk or dawn	39	3.93%	37	3.57%	52	4.55%
	Dark with streetlights	149	15.01%	133	12.85%	124	10.86%
	Dark without streetlights	265	26.69%	279	26.96%	321	28.11%
Weather	Clear	546	54.98%	562	54.30%	613	53.68%
	Cloudy, rain and other weather conditions	447	45.02%	473	45.70%	529	46.32%
Area type	Rural	377	37.97%	420	40.58%	505	44.22%
	Mixed	149	15.01%	138	13.33%	154	13.49%
	Urban	467	47.03%	477	46.09%	483	42.29%
Terrain	Flat	147	14.80%	128	12.37%	127	11.12%

Variable	Attribute	2016		2017		2018	
		Count	Percentage	Count	Percentage	Count	Percentage
	Rolling and mountainous	846	85.20%	907	87.63%	1015	88.88%
Speed limit	<= 35 mph	135	13.60%	122	11.79%	119	10.42%
	40-60 mph	338	34.04%	364	35.17%	406	35.55%
	Above 60 mph	520	52.37%	549	53.04%	617	54.03%
Road classification	Interstate highway	628	63.24%	647	62.51%	718	62.87%
	US highway and State highway	365	36.76%	388	37.49%	424	37.13%
Road configuration	One-way mainline	453	45.62%	456	44.06%	520	45.53%
	Undivided two-way mainline	36	3.63%	40	3.86%	56	4.90%
	Divided two-way mainline with no median barrier	70	7.05%	77	7.44%	77	6.74%
	Divided two-way mainline with median barrier	434	43.71%	462	44.64%	489	42.82%
Collision type	Run-off road	350	35.25%	331	31.98%	335	29.33%
	Overturned	68	6.85%	61	5.89%	68	5.95%
	Hit object	546	54.98%	619	59.81%	718	62.87%
	Other collision type	29	2.92%	24	2.32%	21	1.84%
Crash location	On traffic lanes	552	55.59%	628	60.68%	655	57.36%
	Outside traffic lanes	441	44.41%	407	39.32%	487	42.64%
Ramp type	Off-ramp	511	51.46%	534	51.59%	598	52.36%
	Off-ramp terminal on crossroad	31	3.12%	50	4.83%	51	4.47%
	Merging lane between on-ramp and off-ramp	30	3.02%	28	2.71%	41	3.59%
	On-ramp	375	37.76%	377	36.43%	392	34.33%

Variable	Attribute	2016		2017		2018	
		Count	Percentage	Count	Percentage	Count	Percentage
	On-ramp terminal on crossroad	46	4.63%	46	4.44%	60	5.25%

Table 3-2 Descriptive statistics of the data for multi-vehicle crashes

Variable	Attribute	2016		2017		2018	
		Count	Percentage	Count	Percentage	Count	Percentage
Injury severity	No injury	3216	77.09%	3499	77.91%	3898	79.91%
	Minor injury	928	22.24%	948	21.11%	939	19.25%
	Severe injury	28	0.67%	44	0.98%	41	0.84%
Gender	No female driver involved	1263	30.27%	1365	30.39%	1549	31.75%
	Female driver involved	2909	69.73%	3126	69.61%	3329	68.25%
Alcohol or drugs	Alcohol or drugs not involved	4100	98.27%	4424	98.51%	4806	98.52%
	Alcohol or drugs involved	72	1.73%	67	1.49%	72	1.48%
Speeding violation	Speeding violation not involved	2172	52.06%	2243	49.94%	2499	51.23%
	Speeding violation involved	2000	47.94%	2248	50.06%	2379	48.77%
Inattentiveness	Driver inattentiveness not involved	3678	88.16%	4003	89.13%	4394	90.08%
	Driver inattentiveness involved	494	11.84%	488	10.87%	484	9.92%
Maneuver	Going straight	526	12.61%	635	14.14%	683	14.00%
	Making maneuver action	3646	87.39%	3856	85.86%	4195	86.00%
Car	Car not involved	848	20.33%	955	21.26%	1133	23.23%
	Car involved	3324	79.67%	3536	78.74%	3745	76.77%

Variable	Attribute	2016		2017		2018	
		Count	Percentage	Count	Percentage	Count	Percentage
Sport utility vehicle	Sport utility vehicle not involved	2648	63.47%	2786	62.04%	2991	61.32%
	Sport utility vehicle involved	1524	36.53%	1705	37.96%	1887	38.68%
Pickup	Pickup not involved	3293	78.93%	3586	79.85%	3844	78.80%
	Pickup involved	879	21.07%	905	20.15%	1034	21.20%
Van	Van not involved	3680	88.21%	3950	87.95%	4346	89.09%
	Van involved	492	11.79%	541	12.05%	532	10.91%
Truck	Truck not involved	3754	89.98%	4017	89.45%	4361	89.40%
	Truck involved	418	10.02%	474	10.55%	517	10.60%
Number of vehicles involved	Two vehicles	3900	93.48%	4231	94.21%	4595	94.20%
	More than two vehicles	272	6.52%	260	5.79%	283	5.80%
Road surface condition	Dry	3685	88.33%	3909	87.04%	4018	82.37%
	Wet	447	10.71%	564	12.56%	784	16.07%
	Other road surface conditions	40	0.96%	18	0.40%	76	1.56%
Lighting condition	Daylight	3343	80.13%	3575	79.60%	3880	79.54%
	Dusk or dawn	155	3.72%	172	3.83%	200	4.10%
	Dark with streetlights	345	8.27%	335	7.46%	373	7.65%
	Dark without streetlights	319	7.65%	395	8.80%	414	8.49%
Weather	Clear	3271	78.40%	3468	77.22%	3588	73.55%
	Cloudy, rain and other weather conditions	901	21.60%	1023	22.78%	1290	26.45%
Area type	Rural	1005	24.09%	1120	24.94%	1461	29.95%
	Mixed	597	14.31%	678	15.10%	636	13.04%

Variable	Attribute	2016		2017		2018	
		Count	Percentage	Count	Percentage	Count	Percentage
	Urban	2570	61.60%	2693	59.96%	2781	57.01%
Horizontal alignment	Straight	3432	82.26%	3725	82.94%	4027	82.55%
	Curve	740	17.74%	766	17.06%	851	17.45%
Terrain	Flat	233	5.58%	254	5.66%	282	5.78%
	Rolling and mountainous	3939	94.42%	4237	94.34%	4596	94.22%
Speed limit	<= 35 mph	699	16.75%	725	16.14%	750	15.38%
	40-60 mph	1694	40.60%	1710	38.08%	1974	40.47%
	Above 60 mph	1779	42.64%	2056	45.78%	2154	44.16%
Road classification	Interstate highway	2550	61.12%	2838	63.19%	3053	62.59%
	US highway and State highway	1622	38.88%	1653	36.81%	1825	37.41%
Road configuration	One-way mainline	1876	44.97%	2006	44.67%	2257	46.27%
	Undivided two-way mainline	417	10.00%	449	10.00%	482	9.88%
	Divided two-way mainline with no median barrier	405	9.71%	421	9.37%	427	8.75%
	Divided two-way mainline with median barrier	1474	35.33%	1615	35.96%	1712	35.10%
Collision type	Rear-end collision	2621	62.82%	2843	63.30%	2992	61.34%
	Sideswipe collision	902	21.62%	965	21.49%	1148	23.53%
	Angle collision	485	11.63%	490	10.91%	496	10.17%
	Other collision type	164	3.93%	193	4.30%	242	4.96%
Crash location	On traffic lanes	4130	98.99%	4436	98.78%	4834	99.10%
	Outside traffic lanes	42	1.01%	55	1.22%	44	0.90%

Variable	Attribute	2016		2017		2018	
		Count	Percentage	Count	Percentage	Count	Percentage
Ramp type	Off-ramp	1705	40.87%	1866	41.55%	2138	43.83%
	Off-ramp terminal on crossroad	989	23.71%	1072	23.87%	1044	21.40%
	Merging lane between on-ramp and off-ramp	220	5.27%	233	5.19%	261	5.35%
	On-ramp	1027	24.62%	1052	23.42%	1174	24.07%
	On-ramp terminal on crossroad	231	5.54%	268	5.97%	261	5.35%

3.3 Method

Temporal transferability refers to the applicability of safety assessment model for new data from other temporal units. This study deals with transferability over time, where the models developed are used for safety assessment using the data from different years[3]. Formulations of the proposed temporal unconstrained parameters models and partially temporal constrained parameters models are presented in the following section.

3.3.1 Modeling approach

In this study, random parameters multinomial logit model with heterogeneity in means and variances is applied to measure the association between influencing factors and crash injury severity at ramp areas.

To estimate the probability of crash injury severity level j (no injury, minor injury, and severe injury) of observation n, the injury severity function U_{nj} is given by (***Washington et al., 2020***),

$$U_{nj} = \boldsymbol{\beta}_j{}'\boldsymbol{x}_n + \varepsilon_{nj} \tag{3-1}$$

where $\boldsymbol{\beta}_j$ is a vector of mean coefficients for injury severity level j, $\boldsymbol{x}_n$ is a vector of explanatory variables for observation n, and ε_{nj} is the generalized extreme value distributed error term.

Then, the multinomial logit model can be expressed as,

$$\mathrm{P}_{nj}|\beta_j = \frac{\exp(\boldsymbol{\beta}_j{}'x_n)}{\sum_{j=1}^{J}\exp(\boldsymbol{\beta}_j{}'x_n)} \tag{3-2}$$

[3] Given the difference in crash mechanisms between single-vehicle crashes and multi-vehicles, transferability between single-vehicle crashes and multi-vehicle crashes is out of the scope of this study.

where P_{nj} is the probability of injury severity for crash n. Then the unconditional probability is specified as,

$$\mathrm{P}_{nj} = \int \frac{\exp(\boldsymbol{\beta}_j{}' \boldsymbol{x}_n)}{\sum_{j=1}^{J} \exp(\boldsymbol{\beta}_j{}' \boldsymbol{x}_n)} f(\boldsymbol{\beta}|\boldsymbol{\varphi})\, d\boldsymbol{\beta} \tag{3-3}$$

where $f(\boldsymbol{\beta}|\boldsymbol{\varphi})$ is the probability density function for vector $\boldsymbol{\beta}$, and $\boldsymbol{\varphi}$ is the vector of parameters that defines the probability density function.

The random parameters with heterogeneity in means and variances can be expressed as,

$$\boldsymbol{\beta}_{nj} = \boldsymbol{\beta}_j + \boldsymbol{\Theta}_{nj}\boldsymbol{z}_{nj} + \boldsymbol{\sigma}_{nj}\exp\left(\boldsymbol{\Psi}_{nj}\boldsymbol{w}_{nk}\right)\boldsymbol{\upsilon}_{nj} \tag{3-4}$$

where $\boldsymbol{\beta}_j$ is the vector of mean parameters of all observations defined in **Eq. (3-1)**, $\boldsymbol{\Theta}_{nj}$ is a matrix of estimated parameters, $\boldsymbol{z}_{nj}$ is a vector of explanatory variables that capture heterogeneity in means, $\boldsymbol{w}_{nj}$ is a vector of explanatory variables that capture heterogeneity in the standard deviation $\boldsymbol{\sigma}_{nj}$ with a matrix of parameters $\boldsymbol{\Psi}_{nj}$, and $\boldsymbol{\upsilon}_{nj}$ is a random term, which follows the standard normal distribution.

The simulated maximum likelihood approach with 1,000 Halton draws is used to estimate the parameters (***Train, 2009***). Average marginal effects calculated by averaging the individual observations are also computed to measure the effect of one-unit change in the explanatory variable on the probability of specific injury severity.

3.3.2 Transferability assessment

3.3.2.1 Temporal unconstrained parameters model

First, separated temporal unconstrained parameter models based on the data from different years are developed for variable determination. The likelihood ratio tests are carried out to assess the overall transferability across time of crash injury severity models (***Washington et al., 2020***). A chi-square test statistic that indicates the stability of estimated parameters over two years can be given by,

$$\mathrm{X}^2 = -2\left[LL\left(\beta_{t_2 t_1}\right) - LL\left(\beta_{t_1}\right)\right] \tag{3-5}$$

where $LL(\beta_{t_2t_1})$ is the log-likelihood at the convergence of a model using converged parameters from year t_2 and data from year t_1, and $LL(\beta_{t_1})$ is the log-likelihood at the convergence of the model using data from year t_1. The degree of freedom is equal to the number of estimated parameters in year t_2. This test can be reversed using parameters from year t_1 and data from year t_2 for result comparison. When the chi-square test statistic is significant, the null hypothesis that the model parameters of year t_1 and t_2 are the same can be rejected.

3.3.2.2 Temporal constrained parameters model

Then, transferability over time (e.g., temporal instability and stability) can be considered using temporal constrained and unconstrained modeling approaches. Partially temporal constrained model includes both unconstrained parameters (parameters of variables in separated temporal unconstrained parameter models are different from one year to the next) and constrained parameters (parameters are the same over year). It provides an efficient way to test for temporal shifting parameters of specific variables (***Alnawmasi and Mannering, 2023***; ***Dzinyela et al., 2024***). For the maximum likelihood estimation, a likelihood ratio test is used to compare temporal unconstrained and constrained parameters for each variable as,

$$X^2 = -2[LL(\beta_C) - LL(\beta_U)] \tag{3-6}$$

where $LL(\beta_C)$ is the log-likelihood at convergence of temporally constrained model, and $LL(\beta_U)$ is the log-likelihood at convergence of temporally unconstrained model. Null hypothesis is that the constrained and unconstrained parameters are equal. If the null hypothesis can be rejected, the unconstrained parameters are warranted. Degree of freedom is equal to the difference in the number of estimated parameters.

Then, another likelihood ratio test is used to compare temporally unconstrained and partially constrained models as,

$$X^2 = -2\left[LL(\beta_{PC}) - \sum LL(\beta_{U,T})\right] \tag{3-7}$$

where $LL(\beta_{PC})$ is the log-likelihood at convergence of partially constrained model, and $LL(\beta_{U,T})$ is the log-likelihood at convergence of unconstrained model for year T. Null hypothesis is that the temporal unconstrained and partially constrained parameters are equal, with degree of freedom equal to the difference in the number of estimated parameters. If the null hypothesis is not rejected, the partially constrained model is warranted (For more discussion about partially constrained model, see ***Alnawmasi and Mannering (2023)***).

3.3.2.3 Out-of-sample prediction

Simulation-based approach is adopted for the out-of-sample prediction of proposed random parameters multinomial logit model with heterogeneity in means and variances. This is to resolve the problem of simplification of random parameters (***Hou et al., 2022***). For example, predicted probability of crash injury severity level j based on estimated parameters is given by,

$$P_n(j) = \frac{1}{R}\sum_{r=1}^{R}\frac{\exp\{[\boldsymbol{\beta}_j + \boldsymbol{\Theta}_{nj}\boldsymbol{z}_{nj} + \boldsymbol{\sigma}_{nj}\exp(\boldsymbol{\Psi}_{nj}\boldsymbol{w}_{nk})\boldsymbol{\upsilon}_{nj,r}]\boldsymbol{x}_n\}}{\sum_{j=1}^{J}\exp\{[\boldsymbol{\beta}_j + \boldsymbol{\Theta}_{nj}\boldsymbol{z}_{nj} + \boldsymbol{\sigma}_{nj}\exp(\boldsymbol{\Psi}_{nj}\boldsymbol{w}_{nk})\boldsymbol{\upsilon}_{nj,r}]\boldsymbol{x}_n\}} \tag{3-8}$$

where R is number of draws from a predefined distribution. In this study, 1,000 Halton draws are adopted to provide adequate numerical iterations for precise parameter estimation of the simulation-based model. Differences between the average out-of-sample predicted probabilities and average in-sample predicted probabilities are estimated to assess the transferability of the model.

3.4 Results and discussion

Table 3-3 and **Table 3-4** present the partially constrained parameter estimation results of random parameters logit models with heterogeneity in means and variances for single- and multi-vehicle crashes [4]. Additionally, average marginal effects of

[4] It should be noted that the *t*-statistics in the model estimation results are only suggestive of

explanatory variables for crash injury severity are also given. The temporally unconstrained parameter estimation results are presented in the **Appendix**. The chi-square test statistic for temporally unconstrained and partially constrained models for single-vehicle crashes is 5.95 with 8 degrees of freedom. There is no sufficient evidence to reject the null hypothesis that the temporal unconstrained and partially constrained models are equal. On the other hand, the chi-square test statistic for multi-vehicle crashes is 13.51 with 15 degrees of freedom. Again, there is no sufficient evidence to reject the null hypothesis that the temporal unconstrained and partially constrained models are equal. Therefore, partially constrained models are warranted. Detailed discussion of similarities and differences in the results of parameter estimation between single- and multi-vehicle crashes will be focused on the partially constrained models in the remaining of this section.

3.4.1 Driver characteristics

For the effects of driver characteristics on single-vehicle crashes, as indicated by the marginal effects in **Table 3-3**, probabilities of minor injury and severe injury are higher when the driver is female in year 2017. Likelihood of severe injury in 2016 is also higher when a young driver (under 25 in age) or an older driver (over 59 in age) is involved. In addition, probabilities of minor injury and severe injury are higher in all three years when driving under the influence of alcohol and drug. This is consistent with the findings of previous studies (***Wang et al., 2009***; ***Li et al., 2012***; ***Mergia et al., 2013***; ***Zhang et al., 2018***). Furthermore, aberrant driving behaviors including inattentiveness, oversteer and aggressive driving behavior significantly

---

significance (***Alnawmasi and Mannering, 2023***; ***Islam et al., 2023***). The correct measure is the improvement in the log-likelihood at convergence after the inclusion of a variable of interest using the χ^2 distributed likelihood ratio test, with the null hypothesis that the models with and without the variable of interest are the same being rejected. To this end, a significant variable refers to that the null hypothesis is rejected with over 90% confidence.

affect the crash injury severity. Probabilities of minor injury and severe injury are higher when aberrant driving is involved. On the other hand, the likelihood of severe injury in all years is lower when seat belt is used. This implies the temporal stability for the effect of seat belt on crash injury, with a 0.0249 reduction in the probability of severe injury.

For the effects of driver characteristics on multi-vehicle crashes, as shown in **Table 3-4**, likelihood of no injury in the crash is lower when female driver is involved in 2017 and 2018. As indicated by the marginal effects, the probabilities of minor injury and severe injury are higher when driving under the influence of alcohol and drug. This finding confirms the results of previous work on multi-vehicle crashes (***Wu et al., 2014***). Additionally, likelihood of severe injury is lower in all three years when maneuver action is involved in multi-vehicle crash. This could be because when making a maneuver action, the driver must slow down. Speed reduction, even if marginal, is correlated to the reduction in energy dissipation in the collision, and therefore, likelihood of injury reduces (***Wali et al., 2020***). Different from single-vehicle crash models, indicator variables of aberrant driving behavior are not statistically significant. Aberrant driving behaviors are prone to injury in single vehicle crashes only. This should imply the instability for the influences of aberrant driving behaviors on crash outcome across different crash types.

3.4.2 Vehicle attributes

For the effect of vehicle class on single-vehicle crashes, likelihood of severe injury in car is lower in 2016. In particular, the effect of car on crash injury severity is random, following a normal distribution with the mean of -4.26 and standard deviation of 2.57 in 2016. This implies that in few crashes (4.9%), likelihood of injury in car would increase, resulting in a higher probability of severe injury. However, the effect of vehicle class is not statistically significant in all years. This implies the temporal instability for the influences of vehicle type on crash injury severity.

For multi-vehicle crashes, likelihood of severe injury is lower when two vehicles are involved, compared to three and more vehicles involved. This is consistent with previous studies (***Chang and Chien, 2013***; ***Feng et al., 2016***; ***Tamakloe et al., 2020***). Truck tends to increase the likelihood of severe injury in multi-vehicle crashes but is statistically insignificant in single-vehicle crashes. It could be that large vehicles (e.g., trucks) usually have heavier, higher and more rigid vehicle frames. This can protect the occupants in the truck to some extent, but it would increase the probability of injury to the occupants of other vehicles involved due to the energy dissipation and underride[5], particularly for small passenger cars (***Mannering, 2018***; ***Alogaili and Mannering, 2022***). Variation in the effect of the presence of truck on crash severity across crash types (i.e., single- and multi-vehicle crashes) is considerable. This should justify the inherent difference in the crash mechanism between single-vehicle and multi-vehicle crashes. Nevertheless, it is worth exploring the impact of the differences in vehicle mass and speed between the vehicles involved on the crash severity when more comprehensive information is available in experimental studies.

3.4.3 Environmental conditions

For the effects of environmental characteristics on single-vehicle crashes, probability of severe injury under dry road surface condition is higher in all three years as indicated by the marginal effects. Moreover, effect of dry road surface is random in 2017. There is 17% of observations that likelihood of injury is lower under dry road surface condition. However, for the remaining 83% of observations, more severe injuries are more likely to be sustained. The higher probability of severe injury may unveil the effect of drivers' risk compensating behavior under seemingly favorable road conditions (***Fountas et al., 2021***). The likelihood of severe injury is positively

[5] An underride collision occurs when the primary structural components of the colliding vehicles have a height mismatch, causing the vehicle with the lower height to be forcefully wedged beneath the structure of the other vehicle (***Boggess et al., 2010***).

associated with unlighted darkness conditions in 2018. In addition, single vehicle crashes in the rural areas tend to be less severe in 2017, with a 0.0004 lower probability of severe injury and a 0.0037 lower probability of minor injury.

For multi-vehicle crashes, those crashes in poor lighting conditions, including dusk or dawn, dark without streetlights and dark with streetlights have a higher probability of resulting in more severe injuries in all three years as indicated by the marginal effects. Different from single-vehicle crash models, crashes in the rural areas tend to be more severe in 2016 and 2018. This could be attributed to the higher speed limit, fewer traffic control and poor road maintenance in the rural areas (***Geedipally et al., 2014***). Furthermore, there may be a lack of protective devices like shoulder and barrier and healthcare services in the rural areas. This may result in higher risk of severe injuries of road victims (***Lee et al., 2018***).

3.4.4 Roadway design

For the effects of roadway characteristics on single-vehicle crashes, probability of no injury is higher and probability of minor injury is lower for flat terrain, compared to rolling and mountainous terrain in 2017 and 2018. A possible explanation is that flat terrain provides a more favorable road environment for the drivers, allowing vehicles to be kept stable when entering or exiting the ramp areas. In rolling or mountainous terrains, the combined effects of curves and elevation changes can hinder the vehicle stability (***Tulu et al., 2015***). There is a negative association between interstate highway crash and likelihood of no injury in 2016 with a 0.0111 reduction in the probability of minor injury and a 0.0006 reduction in the probability of severe injury.

For multi-vehicle crashes, crashes on interstate highways tend to be less severe in 2016 only with a 0.0049 lower probability of minor injury and a 0.0001 lower probability of severe injury. Effect of higher speed limit (above 60 mph) is random, following a normal distribution with the mean of -1.32 and standard deviation of 1.53

for the likelihood of no injury in 2016, the mean of -1.70 and standard deviation of 2.39 for the likelihood of no injury in 2017, and the mean of 0.72 and standard deviation of 1.40 for the likelihood of minor injury in 2018. In addition, crashes at one-way mainline have lower probabilities of minor and severe injuries in all years.

3.4.5 Crash circumstances

For the effects of crash circumstances on single-vehicle crashes, probabilities of minor injury and severe injury of overturned crash are higher in all years. This is consistent with the findings of previous studies (***Fountas and Anastasopoulos, 2017***; ***Fountas et al., 2018b***; ***Islam et al., 2020***). For the effect of crash location, probabilities of minor injury and severe injury when on the traffic lanes are lower in 2018. Furthermore, probability of minor injury of crash at off-ramp area or on-ramp area is lower in 2016 and 2017 compared to at merging lane areas between on-ramp and off-ramp.

Estimation results for multi-vehicle crashes indicate that probabilities of minor injury and severe injury of sideswipe collision are lower, whereas angle collision tends to be more severe compared to rear-end collision and other crash types in all three years. This is consistent with the findings of previous studies (***Wang et al., 2009***; ***Li et al., 2012***; ***Mergia et al., 2013***; ***Zhang et al., 2018***). It may be because angle collisions usually involve certain degree of side impact. Considering the vehicle structure and safety protection devices, crumple zones on the sides of vehicle could be limited for energy dissipation in the side impact (***O'Neill, 2009***). Additionally, a side impact can cause the occupant’s head to strike the window or door, resulting in head, neck, or back injuries. Moreover, likelihood of minor injury of crash at on-ramp is higher.

3.4.6 Heterogeneity in means and variances

Heterogeneity in means and variances of random parameters are also considered in single- and multi-vehicle crash models. For the likelihood of severe injury of single-vehicle crashes in 2016, effect of car is positively associated with aggressive driving behavior. This finding suggests that the mean of random parameter for car will be modified by other exogenous variables. For example, although likelihood of severe injury of car is lower in 2016, interaction effect with aggressive driving indicates that the likelihood of injury of car could be positively associated with aggressive driving. Furthermore, mean of the random parameter of dry road surface on the likelihood of injury increases with merging lane between on-ramp and off-ramp in 2017. This may be because of risk-compensating behavior when driving under favorable conditions (***Mannering and Bhat, 2014***). For instance, on a dry road, drivers are generally more confident and might travel at a higher speed at the weaving area. The mean of random parameter of rural area increases when the safety belt is used. Regarding the heterogeneity in variances of random parameters of single-vehicle crashes, variance of random parameter of rural area increases for no injury of crash when the driver is female but decreases at higher speed limit area (above 60 mph) in 2018.

In the multi-vehicle crash models, they also have statistically significant heterogeneity in the mean of random parameter. For instance, in 2016, the mean of random parameter of higher speed limit for the likelihood of no injury is higher when the crash is two vehicles involved, or under clear weather condition. In 2017, the mean of random parameter of higher speed limit for the likelihood of no injury decreases when truck is involved in multi-vehicle crash, but increases when the crash is two vehicles involved or on traffic lanes. In contrast, in the multi-vehicle model of year 2018, truck involved crash is positively associated with the mean of random parameter of higher speed limit for minor injury. With regards to the heterogeneity in variances, making maneuver action is negatively associated with the variance of random parameter of higher speed limit for the likelihood of no injury in 2017. In 2018, the variance of random parameter of multi-vehicle crash at higher speed limit area for the likelihood of minor injury is negatively associated with straight horizontal alignment but positively associated with rear-end collision.

Table 3-3 Results of partially constrained parameter estimation for single-vehicle crashes

Variable	Coefficient	t-statistic	Marginal effects		
			No injury	Minor injury	Severe injury
Constant [NI] [2016]	0.50	2.30			
Constant [NI] [2017]	1.43	12.18			
Constant [NI] [2018]	1.05	6.07			
Constant [SI] [2016]	-2.73	-2.54			
Constant [SI] [2017]	-2.31	-2.70			
Constant [SI] [2018]	-1.54	-3.37			
Random parameter (normally distributed)					
Vehicle type					
Car [SI] [2016]	-4.26	-1.48	-0.0008	-0.0002	0.0010
Standard deviation	2.57	1.66			
Road surface condition					
Dry [SI] [2017]	0.80	1.66	-0.0060	-0.0024	0.0084
Standard deviation	0.85	1.50			
Area type					
Rural [NI] [2018]	-1.86	-2.00	0.0041	-0.0041	0.0001
Standard deviation	2.00	1.99			
Heterogeneity in the mean of random parameter					
Car [SI] [2016]: Aggressive driving	3.23	2.03			
Dry [SI] [2017]: Merging lane between on-ramp and off-ramp	2.17	1.91			
Rural [NI] [2018]: Used safety belt	3.06	2.70			
Heterogeneity in the variance of random parameter					
Rural [NI] [2018]: Female driver	0.73	1.69			
Rural [NI] [2018]: Speed limit above 60 mph	-0.60	-1.48			
Gender					
Female [NI] [2017]	-0.37	-2.42	-0.0082	0.0076	0.0006
Age					
Below 25 [SI] [2016]	1.32	1.77	-0.0016	-0.0008	0.0024
Above 59 [SI] [2016]	1.65	1.99	-0.0010	-0.0005	0.0015
Alcohol or drugs					
Driving under the influence of alcohol or drugs [NI] [2016, 2017, 2018]	-0.74	-5.03	-0.0128	0.0116	0.0012
Aberrant driving behavior					

Variable	Coefficient	t-statistic	Marginal effects		
			No injury	Minor injury	Severe injury
Oversteer [NI] [2016, 2017]	-0.47	-2.84	-0.0065	0.0061	0.0004
Inattentiveness [NI] [2016]	-0.61	-2.45	-0.0035	0.0033	0.0002
Aggressive driving [NI] [2017, 2018]	-0.92	-4.04	-0.0074	0.0067	0.0007
Safety belt					
Used [SI] [2016, 2017, 2018]	-2.39	-6.21	0.0166	0.0083	-0.0249
Road surface condition					
Dry [SI] [2016]	2.10	1.85	-0.0054	-0.0031	0.0085
Dry [NI] [2018]	-0.53	-3.17	-0.0182	0.0172	0.0010
Lighting condition					
Dark without streetlights [SI] [2018]	1.38	2.80	-0.0021	-0.0021	0.0042
Area type					
Rural [NI] [2017]	0.18	1.35	0.0041	-0.0037	-0.0004
Terrain					
Flat [MI] [2017]	-0.48	-1.88	0.0026	-0.0027	0.0001
Flat [NI] [2018]	0.64	1.96	0.0029	-0.0027	-0.0002
Road classification					
Interstate highway [NI] [2016]	0.32	2.16	0.0116	-0.0111	-0.0006
Collision type					
Overturned [NI] [2016, 2017, 2018]	-1.40	-8.14	-0.0187	0.0174	0.0013
Crash location					
On traffic lanes [NI] [2018]	0.42	2.51	0.0140	-0.0133	-0.0006
Ramp type					
Off-ramp [NI] [2016]	0.46	2.01	0.0146	-0.0139	-0.0007
Off-ramp [SI] [2017]	1.46	2.46	-0.0052	-0.0021	0.0073
On-ramp [NI] [2016]	0.79	3.32	0.0160	-0.0153	-0.0007
Model statistics					
Number of observations	3170				
Degree of freedom	37				
Log-likelihood at zero (LL(0))	-3482.6010				
Log-likelihood at convergence (LL(β))	-1931.1594				
McFadden R^2	0.4455				

Parameter defined for: [NI] No injury; [MI] Minor Injury; [SI] Severe Injury

Table 3-4 Results of partially constrained parameter estimation for multi-vehicle crashes

Variable	Coefficient	t-statistic	Marginal effects		
			No injury	Minor injury	Severe injury
Constant [NI] [2016]	1.04	10.35			
Constant [NI] [2017]	1.37	16.33			
Constant [NI] [2018]	1.46	18.15			
Constant [SI] [2016]	-3.67	-9.65			
Constant [SI] [2017]	-1.69	-4.29			
Constant [SI] [2018]	-1.75	-4.00			
Random parameter (normally distributed)					
Lighting Condition					
Dark with streetlights [NI] [2017]	-2.07	-1.71	-0.0014	0.0014	0.0000
Standard deviation	3.47	1.84			
Speed Limit					
Above 60 mph [NI] [2016]	-1.32	-4.16	-0.0030	0.0030	0.0001
Standard deviation	1.53	2.18			
Above 60 mph [NI] [2017]	-1.70	-2.54	-0.0040	0.0038	0.0002
Standard deviation	2.39	3.27			
Above 60 mph [MI] [2018]	0.72	2.66	-0.0023	0.0024	-0.0001
Standard deviation	1.40	1.89			
Heterogeneity in the mean of random parameter					
Dark with streetlights [NI] [2017]: Two vehicles	3.61	1.72			
Above 60 mph [NI] [2016]: Involvement of alcohol or drugs for drivers	-1.78	-2.77			
Above 60 mph [NI] [2016]: Two vehicles	1.93	3.78			
Above 60 mph [NI] [2016]: Clear weather	0.43	1.95			
Above 60 mph [NI] [2017]: Truck involved	-0.47	-1.76			
Above 60 mph [NI] [2017]: Two vehicles	1.59	3.52			
Above 60 mph [NI] [2017]: On traffic lanes	1.12	1.77			
Above 60 mph [MI] [2018]: Truck involved	0.67	2.70			
Above 60 mph [MI] [2018]: Two vehicles	-1.76	-5.11			
Heterogeneity in the variance of random parameter					
Above 60 mph [NI] [2017]: Making maneuver action	-0.38	-1.78			
Above 60 mph [MI] [2018]: Straight horizontal alignment	-0.68	-2.62			
Above 60 mph [MI] [2018]: Rear-end collision	0.83	1.90			

Variable	Coefficient	t-statistic	Marginal effects		
			No injury	Minor injury	Severe injury
Gender					
Female driver involved [MI] [2016]	0.16	1.75	-0.0052	0.0052	0.0001
Female driver involved [NI] [2017, 2018]	-0.42	-6.11	-0.0286	0.0278	0.0008
Alcohol or drugs					
Involvement of alcohol or drugs for drivers [NI] [2017, 2018]	-1.37	-5.87	-0.0025	0.0024	0.0001
Maneuver					
Making maneuver action [SI] [2016, 2017, 2018]	-1.17	-4.27	0.0026	0.0012	-0.0038
Truck					
Truck involved [SI] [2016, 2017, 2018]	1.74	6.56	-0.0020	-0.0010	0.0030
Number of vehicles involved					
Two vehicles [SI] [2017, 2018]	-1.53	-4.54	0.0032	0.0013	-0.0045
Lighting condition					
Dusk or dawn [SI] [2016]	1.67	2.09	-0.0002	-0.0001	0.0002
Dusk or dawn [NI] [2017]	-0.54	-2.71	-0.0012	0.0011	0.0001
Dark with streetlights [NI] [2016]	-0.45	-3.32	-0.0021	0.0021	0.0001
Dark without streetlights [NI] [2018]	-0.28	-1.88	-0.0012	0.0012	0.0001
Area type					
Rural [NI] [2016, 2018]	-0.23	-3.27	-0.0059	0.0057	0.0002
Road classification					
Interstate highway [NI] [2016]	0.21	2.25	0.0050	-0.0049	-0.0001
Road configuration					
One-way mainline [NI] [2016, 2017, 2018]	0.43	7.77	0.0230	-0.0225	-0.0005
Collision type					
Sideswipe collision [NI] [2016, 2017, 2018]	1.00	12.51	0.0193	-0.0186	-0.0007
Angle collision [NI] [2016, 2017, 2018]	-0.53	-7.58	-0.0125	0.0122	0.0003
Ramp type					
On-ramp [MI] [2016, 2017]	0.19	2.67	-0.0042	0.0043	-0.0001
Model statistics					
Number of observations	13541				
Degree of freedom	42				
Log-likelihood at zero (LL(0))	-14876.3090				
Log-likelihood at convergence (LL(β))	-6860.3030				
McFadden R^2	0.5388				

Parameter defined for: [NI] No injury; [MI] Minor Injury; [SI] Severe Injury

3.5 Transferability assessment

Table 3-5 and **Table 3-6** show the results of likelihood ratio tests for temporal stability of single- and multi-vehicle crashes based on temporally unconstrained models. Additionally, **Table 3-7** and **Table 3-8** show the differences between out-of-sample predicted probabilities and in-sample predicted probabilities for single- and multi-vehicle crashes respectively.

3.5.1 Temporal stability

The results of likelihood ratio tests for single-vehicle crashes are shown in **Table 3-5**. For example, null hypothesis (effects of influencing factors on crash injury severity are consistent across years) of the model using converged parameters of 2017 model and 2016 data (X^2 =32.834 with 14 degrees of freedom) can be rejected at the 1% level of significance. This justifies the statistically significant variations and the existence of possible temporal instability. In contrast, there is no sufficient evidence to reject the null hypothesis of model using converged parameters of 2018 model and 2016 data, parameters of 2016 model and 2017 data, parameters of 2018 model and 2017 data. Similarly, for multi-vehicle crashes, as shown in **Table 3-6**, there is no sufficient evidence (24% confidence) to reject the null hypothesis of the model using converged parameters of 2017 model and 2018 data, parameters of 2018 model and 2017 data, and parameters of 2017 model and 2018 data. Some likelihood ratio tests results exhibit significant temporal inconsistency in the effects of influencing factors on crash injury severity, while others do not.

Table 3-5 Results of likelihood ratio tests for temporal stability of single-vehicle crash

	t_2		
t_1	**2016**	**2017**	**2018**
2016	-	32.834 (14) [< 0.01]	22.648 (15) [0.09]
2017	23.417 (16)	-	19.040 (15)

	t_2		
t_1	**2016**	**2017**	**2018**
	[0.10]		[0.21]
2018	56.508 (16) [< 0.01]	42.766 (14) [< 0.01]	-

Note: Degrees of freedom in parentheses and level of significance in brackets.

Table 3-6 Results of likelihood ratio tests for temporal stability of multi-vehicle crash

	t_2		
t_1	**2016**	**2017**	**2018**
2016	-	10.843 (21) [0.97]	12.462 (18) [0.82]
2017	55.721 (18) [< 0.01]	-	27.959 (18) [0.06]
2018	34.194 (18) [0.01]	16.162 (21) [0.76]	-

Note: Degrees of freedom in parentheses and level of significance in brackets.

3.5.2 Out-of-sample prediction

For single vehicle crashes, as shown in **Table 3-7**, differences in the average predicted probabilities between 2016 and 2017 (using estimated parameters of 2016 to predict outcome probabilities with crash data of 2017 and using estimated parameters of 2017 to predict outcome probabilities with crash data of 2017) for no injury, minor injury, and severe injury are -0.0318, +0.0348, and -0.0030, respectively. This implies, for example, that if the parameters of 2016 did not shift in 2017, the probability of minor injury in 2017 would have been higher. In other words, the parameter shifts observed from 2016 to 2017 resulted in a lower probability of minor injury. Additionally, similar results are observed between 2016 and 2018, with which the differences for no injury, minor injury, and severe injury are -0.0182, +0.0181, and 0.0001, respectively. In contrast, opposite results are observed between 2017 and 2018, with which the differences for no injury, minor injury, and severe injury are +0.0164, -0.0222, and

+0.0058, respectively. This means that the aggregate effect of parameter shifts from 2017 to 2018 has resulted in a higher probability of no injury and severe injury but a lower probability of minor injury. For multi-vehicle crashes, as shown in **Table 3-8**, differences in the average predicted probabilities between 2016 and 2017 for no injury, minor injury, and severe injury are -0.0036, +0.0068, and -0.0033, respectively. Additionally, differences between 2016 and 2018 for no injury, minor injury, and severe injury are -0.0233, +0.0255, and -0.0021, respectively. However, when estimated parameters of 2017 are used to predict outcome probabilities with crash data of 2018, differences for no injury, minor injury, and severe injury are -0.0148, +0.0140, and +0.0008 respectively. Reasons for these observed shifts could be changes in road infrastructure and vehicle safety features (***Chen et al., 2017***), as well as changes in driver behavior and traffic law enforcement (***Yasmin et al., 2022***). Such findings underscore the importance of accounting for temporal instability in the crash severity models (***Mannering, 2018***). However, it is noteworthy that these are transferability assessments of the overall performance of the models. Even if the model's overall performance does not always exhibit temporal stability and transferability, the effects of some factors on the probability of injury severity remain stable over time. More detailed discussions will be provided in the next section.

Table 3-7 Difference in the average predicted probabilities for single vehicle crash

Base year	Prediction year					
	2017			2018		
	No injury	Minor injury	Severe injury	No injury	Minor injury	Severe injury
2016	-0.0318	0.0348	-0.0030	-0.0182	0.0181	0.0001
2017	-	-	-	0.0164	-0.0222	0.0058

Table 3-8 Difference in the average predicted probabilities for multi-vehicle crash

Base year	Prediction year					
	2017			2018		
	No injury	Minor injury	Severe injury	No injury	Minor injury	Severe injury
2016	-0.0036	0.0068	-0.0033	-0.0233	0.0255	-0.0021
2017	-	-	-	-0.0148	0.0140	0.0008

3.6 Concluding remarks

Highway ramp areas are prone to crash and severe injury. This study on temporal transferability assessment of crash injury severity models at ramp areas contributes to the field by addressing the issues regarding the accuracy and reliability of crash injury severity models. Several factors that affect the crash injury severity of single-vehicle and multi-vehicle crashes at ramp areas are considered. Random parameters multinomial logit regression model with heterogeneity in means and variances is adopted to measure the association between possible influencing factors and crash severity at ramp areas based on the crash data from the North Carolina State of the United States in 2016-2018, with which the effects of unobserved heterogeneity and temporal instability are considered. Factors including driver characteristics, vehicle attributes, environmental conditions, roadway design, and crash circumstances are considered. Results indicate that there are considerable differences for the effects of aberrant driving, vehicle type, area type and crash location on the likelihood of injury between single-vehicle and multi-vehicle crashes. For example, truck involvement tends to increase the likelihood of injury in multi-vehicle crash but is statistically insignificant in single-vehicle crash. Additionally, there are opposite effects for the crashes in rural areas on the likelihood of injury between single-vehicle and multi-vehicle crashes. This justifies the need of developing and implementing targeted traffic control and management strategies that can reduce the risk of single-vehicle and multi-vehicle crashes separately. In particular, it is vital to implement traffic control measures like speed enforcement cameras and variable message signs to deter against the aberrant driving behavior at the ramp areas. Therefore, risk of single-vehicle crash can be mitigated. On the other hand, it is worth exploring the effectiveness of advanced driver assistant system that can mitigate the risk of multi-vehicle crashes involving trucks and other heavy vehicles at the ramp areas. Furthermore, partially constrained modeling approach provides an efficient way to test for temporally shifting parameters by combining all data and defining parameters for each period. Results of partially constrained model and transferability assessments indicate that remarkable temporal stability and instability coexist. From the perspective of decision-makers, it is crucial to pay attention to both time-constant and time-varying variables when they significantly affect crash injury severity. Time-constant variables provide a consistent baseline for risk assessment and control, ensuring that ongoing safety measures are effective. On the other hand, parameters that change over time are important for identifying new challenges, as they can indicate emerging threats or the need for changes in prevailing policy strategies. By exploring the temporal transferability of crash injury severity model, understanding the shifts in the effects of significant factors on the crash outcome could be enhanced.

Chapter 4 Correcting for Endogeneity of Crash Type in Crash Injury Severity at Highway Ramp Areas

4.1 Introduction

Studies have examined the association between crash injury risk at ramp areas and influential factors including roadway design, environmental and weather conditions, traffic flow characteristics, driver factors, and crash circumstances. For example, there is a significant association between crash type and injury severity for the crashes at ramps (***Mergia et al., 2013***; ***Song et al., 2024b***). Regarding single-vehicle crashes at ramp areas, the likelihood of severe injury of rollover crashes is higher than that of hit-object crashes (***Mergia et al., 2013***). This suggests that remedial roadway design and traffic control measures including re-alignment, road barrier, shoulder lane, and warning sign could help mitigate the risk of specific crash types at ramp areas. In previous studies, crash type was typically incorporated into the crash severity model as an independent input variable. However, as is the case in any crash data, there could exist unobserved factors, including driver behavior and vehicle performance, that may affect both the crash type and crash severity simultaneously. For example, reckless driving behavior like speeding and making a sharp turn can increase the chance of loss of control, resulting in rollover crash. At the same time, such behavior is often associated with a higher likelihood of severe injury. Furthermore, the likelihood of certain crash types is often lower for the vehicles that are equipped with advanced safety features like electronic stability control, and anti-lock braking systems. Such safety features can also reduce the injury risk in a crash. Hence, the effect of crash type on crash injury severity could be endogenous.

Unfortunately, it was observed that the endogenous effects of potential risk factors at highway ramp areas are rarely considered in conventional road safety studies. This can result in bias of parameter estimation and misinterpretation of the effects of influencing factors on the crash risk (***Mannering and Bhat, 2014***; ***Mannering et al., 2020***). To this end, it is important to adopt efficient analytic methods that can identify and measure such endogenous effects in the analysis of road safety.

The objective of this paper is to explore the endogenous effect of crash type on injury severity of single-vehicle crash at ramp area by developing a random parameters recursive bivariate probit model that can adequately capture the endogeneity and quantify its effect on crash injury severity. In the proposed simultaneous model, crash type is regarded as the treatment variable. Additionally, the effect of unobserved heterogeneity is considered using random parameters specification with heterogeneity in the means.

The remainder of this chapter is structured as follows. **Section 4.2** presents the data used, and analysis methods are given in **Section 4.3**. **Section 4.4** presents the results of parameter estimation. Finally, the concluding remarks are given in **Section 4.5**.

4.2 Data

This study focuses on the injury severity among single-vehicle crashes at ramp areas in the state of North Carolina in a three-year period (2016-2018)[6]. The source of the data is the Federal Highway Administration (FHWA) Highway Safety Information System (HSIS). In the HSIS, crash data are stratified into five classes: fatal injury, suspected serious injury, suspected minor injury, possible injury, and no injury. To avoid the problem of data imbalance (often manifest as relatively sparse records of

[6] It is important to consider the varying effects of exogenous variables across years. However, the major focus of this study is the endogeneity of crash type on injury severity. Addressing temporal instability and out-of-sample prediction is beyond the study scope in this chapter.

severe injury crashes), four injury classes were combined into one class as "injury"[7]. For each crash, the data contained information on driver characteristics, vehicle attributes, environmental conditions, roadway design, and crash circumstance. The total number of observations is 3,170 (one observation for each crash). The descriptive statistics of the sample data are summarized in **Table 4-1**.

For single-vehicle crashes, common crash types are rollover and hit-object crashes (***NHTSA, 2024***). A rollover crash involves a vehicle rotating at least one quarter of a revolution and ending up on its side or end. A hit-object crash involves a vehicle hitting a fixed or non-fixed object. In this study, crash type is dichotomous: rollover or hit-object. As shown in **Table 4-1**, 6.2% of the sample are rollover crashes. The distribution of the sample with respect to injury severity and crash type is given in **Table 4-2**. The result of a chi-square test indicates that the null hypothesis that the injury severities of single-vehicle crashes among different crash types are the same can be rejected at the 1% level of significance (critical chi-squared value of 80.57 with 1 degree of freedom). Also, it was found that no injury crashes (74.2%) were predominant among the hit-object crashes. In addition, there is possibly exists an endogenous effect of crash type on crash injury severity. To this end, crash type is considered as an explanatory variable in the crash injury severity model, and the outcome variable in the crash type model in the simultaneous equation system. Therefore, influences of crash types on crash injury severity can be estimated, and the possible endogenous effects can be accounted for.

[7] In line with previous injury severity analyses of single-vehicle crashes (***Behnood and Mannering, 2019***; ***Fountas et al., 2020***; ***Song et al., 2024b***), the injury severity outcome of a crash would be determined based on the victim(s) who suffers from the most severe injury.

Table 4-1 Descriptive statistics of the data

Variable	Attribute	Count	Percentage
Dependent Variables			
Injury severity	No injury	2294	70.80%
	Injury	876	27.39%
Crash type	Hit-object crash	2973	93.79%
	Rollover crash	197	6.21%
Driver Characteristics			
Gender	Male	2006	63.28%
	Female	1164	36.72%
Age	Below 25	1085	34.23%
	25-39	1101	34.73%
	40-59	722	22.78%
	60 or above	262	8.26%
Alcohol or Drugs	Not under the influence of alcohol or drugs	2915	91.96%
	Driving under the influence of alcohol or drugs	255	8.04%
Aberrant driving behavior	Speeding violation	1148	36.21%
	Oversteer	315	9.94%
	Inattentiveness	242	7.63%
	Aggressive driving	177	5.58%
	Other aberrant driving behavior	843	26.59%
	No violation driving behavior	445	14.04%
Safety belt	Not used	202	6.37%
	Used	2968	93.63%
Maneuver	Changing lanes	498	15.71%
	Going straight	2250	70.98%
	Making turn	285	8.99%
	Other maneuver action	137	4.32%
Vehicle Attributes			
Vehicle type	Car	1953	61.61%
	Sport utility vehicle	548	17.29%
	Pickup	354	11.17%
	Van	102	3.22%
	Truck	180	5.68%
	Other vehicle types	33	1.04%
Environmental Conditions			
Road surface condition	Dry	1818	57.35%
	Wet	1169	36.88%

Variable	Attribute	Count	Percentage
	Other road surface conditions	183	5.77%
Lighting condition	Daylight	1761	55.55%
	Dusk or dawn	128	4.04%
	Dark with streetlights	406	12.81%
	Dark without streetlights	875	27.60%
Weather	Clear	1721	54.29%
	Cloudy, rain and other weather conditions	1449	45.71%
Area type	Rural	1302	41.07%
	Mixed	441	13.91%
	Urban	1427	45.02%
Roadway Design			
Horizontal alignment	Straight	1599	50.44%
	Curve	1571	49.56%
Speed limit	<= 35 mph	376	11.86%
	40-60 mph	1108	34.95%
	Above 60 mph	1686	53.19%
Road classification	Interstate highway	1993	62.87%
	US highway and State highway	1177	37.13%
Road configuration	One-way mainline	1429	45.08%
	Undivided two-way mainline	132	4.16%
	Divided two-way mainline with no median barrier	224	7.07%
	Divided two-way mainline with median barrier	1385	43.69%
Crash Circumstances			
Crash location	On traffic lanes	1835	57.89%
	Outside traffic lanes (on road shoulder)	1335	42.11%
Ramp type	Off-ramp	1775	55.99%
	Merging lane between on-ramp and off-ramp	99	3.12%
	On-ramp	1296	40.88%

Table 4-2 Cross tabulation of injury severity and crash type

Injury severity	Crash type		Total
	Rollover crash	Hit-object crash	
No injury	88 (44.67%)	2206 (74.20%)	2294 (72.37%)
Injury	109 (55.33%)	767 (25.80%)	876 (27.63%)
Overall	197 (100.00%)	2973 (100.00%)	3170 (100.00%)

Note: The numbers in the parentheses represent column percentages.

4.3 Method

In this study, the outcome variables are dichotomous in both crash type and injury severity models. To measure the association between influencing factors and crash injury severity considering the endogeneity of crash type, a random parameters recursive bivariate probit model with heterogeneity in the means and variances is proposed. Let i (i = 1, 2, . . ., I) be an index that represents a crash observation. The general specification of a recursive simultaneous equations system for the joint crash type and injury severity model is given by (***Hensher et al., 2015***; ***Greene, 2018***; ***Washington et al., 2020***),

$$\begin{gathered} r_i^* = \boldsymbol{\beta}_1{}'\boldsymbol{z}_i + \varepsilon_{i,1} \\ r_i = 1, \text{if } r_i^* > 0, r_i = 0 \ otherwise \\ y_i^* = \boldsymbol{\beta}_2{}'\boldsymbol{x}_i + \gamma r_i + \varepsilon_{i,2} \\ y_i = 1, \text{if } y_i^* > 0, y_i = 0 \ otherwise \end{gathered} \tag{1}$$

where r_i is a binary indicator of crash type, y_i is a binary indicator of injury severity, $\boldsymbol{z}_i$ and $\boldsymbol{x}_i$ are column vectors of explanatory variables, $\boldsymbol{\beta}_1{}'$ and $\boldsymbol{\beta}_2{}'$ are the corresponding row vectors of estimated parameters, $\varepsilon_{i,1}$ and $\varepsilon_{i,2}$ are the error terms, respectively.

Given the explanatory variables, the joint probability for $y_i = 1, r_i = 1$ is written as:

$$\text{Prob}(y_i = 1, r_i = 1 | \boldsymbol{x}_i, \boldsymbol{z}_i) = \text{Prob}(y_i^* > 0, r_i^* > 0) \tag{2}$$

When $\varepsilon_{i,1}$ and $\varepsilon_{i,2}$ are distributed as bivariate standard normal with correlation ρ[8] as,

---

[8] Note that the interpretation of the sign of the correlation parameter ρ in recursive bivariate probit model is not the same as that for the bivariate probit model. Herein, the significant ρ just indicates that the binary dependent variable in one equation is an endogenous independent variable in the other equation. When the effect of the endogenous variable is taken into account, the correlation between the errors terms is not necessarily of the same sign

$$\begin{pmatrix} \varepsilon_{i,1} \\ \varepsilon_{i,2} \end{pmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right] \tag{3}$$

The joint probability density function of recursive bivariate probit model for $y_i = 1, r_i = 1$ can be expressed as,

$$\text{Prob}(y_i = 1, r_i = 1 | \boldsymbol{x}_i, \boldsymbol{z}_i) = \boldsymbol{\Phi}_2(\boldsymbol{\beta}_2{}'\boldsymbol{x}_i + \gamma r_i, \boldsymbol{\beta}_1{}'\boldsymbol{z}_i, \rho). \tag{4}$$

There are four cases for this bivariate probit model, with the other three joint probabilities given by,

$$\begin{aligned} \text{Prob}(y_i = 1, r_i = 0 | \boldsymbol{x}_i, \boldsymbol{z}_i) &= \boldsymbol{\Phi}_2(\boldsymbol{\beta}_2{}'\boldsymbol{x}_i + \gamma r_i, -\boldsymbol{\beta}_1{}'\boldsymbol{z}_i, -\rho), \\ \text{Prob}(y_i = 0, r_i = 1 | \boldsymbol{x}_i, \boldsymbol{z}_i) &= \boldsymbol{\Phi}_2[-(\boldsymbol{\beta}_2{}'\boldsymbol{x}_i + \gamma r_i), \boldsymbol{\beta}_1{}'\boldsymbol{z}_i, -\rho], \\ \text{Prob}(y_i = 0, r_i = 0 | \boldsymbol{x}_i, \boldsymbol{z}_i) &= \boldsymbol{\Phi}_2[-(\boldsymbol{\beta}_2{}'\boldsymbol{x}_i + \gamma r_i), -\boldsymbol{\beta}_1{}'\boldsymbol{z}_i, \rho]. \end{aligned} \tag{5}$$

Hence, the log-likelihood function for the proposed model is,

$$\ln L = \sum_{i=1}^{n} \ln \boldsymbol{\Phi}_2\left(q_{y,i}(\boldsymbol{\beta}_2{}'\boldsymbol{x}_i + \gamma r_i), q_{r,i}(\boldsymbol{\beta}_1{}'\boldsymbol{z}_i), q_{y,i}q_{r,i}\rho\right) \tag{6}$$

where $q_{y,i} = 2y_i - 1$ and $q_{r,i} = 2r_i - 1$, $\boldsymbol{\Phi}_2(.)$ denotes the bivariate standard normal cumulative distribution function.

Furthermore, accounting for unobserved heterogeneity, the random parameters model with heterogeneity in mean and variance is expressed as (***Greene et al., 2006; Washington et al., 2020***),

$$\boldsymbol{\beta}_i = \boldsymbol{\beta} + \boldsymbol{\Theta}\boldsymbol{z}_i + \boldsymbol{\Sigma}_{\boldsymbol{i}}^{\boldsymbol{1/2}}\boldsymbol{v}_i \tag{7}$$

---

as the endogenous relationship. That is, the correlation parameter in recursive model merely indicates the presence of endogeneity, rather than interpreting the correlation between outcome variables. For more discussion about the different interpretation of the correlation parameters in the bivariate probit model and the recursive bivariate probit model, see ***Filippini et al. (2018)***.

Let $\boldsymbol{\Sigma}_i^{1/2} = \text{Diag}[\sigma_{i1}, \sigma_{i2}, \ldots \sigma_{ik}]$, then

$$\sigma_{ik} = \sigma_k \times \exp(\boldsymbol{\psi}_k' \boldsymbol{w}_i) \tag{8}$$

where $\boldsymbol{\beta}$ is the vector of mean coefficients for all observations defined in Eq. **(1)**, $\boldsymbol{\Theta}$ is a matrix of estimated parameters, $\boldsymbol{z}_i$ is a vector of explanatory variables that capture heterogeneity in the means of random parameters, the scale factor σ_{ik} which provides the standard deviation of the kth random parameter is then arrayed on the diagonal of the diagonal matrix $\boldsymbol{\Sigma}_i^{1/2}$, $\boldsymbol{w}_i$ is a vector of explanatory variables that capture heterogeneity in the variances[9], $\boldsymbol{\psi}_k'$ is the kth row elements of the matrix of estimated parameters $\boldsymbol{\Psi}$, and $\boldsymbol{v}_i$ is a primitive random vector, which follows the standard normal distribution.

The simulated maximum likelihood approach with 1,000 Halton draws is used to estimate the parameters. Lastly, to estimate the change in probability of the injury severity because of the change in $\boldsymbol{x}$ and/or $\boldsymbol{z}$, the marginal effects[10] can be estimated by (***Greene, 1998***, ***2018***),

$$\begin{aligned}
ME &= E[y|\boldsymbol{x}, \boldsymbol{z}, X = 1] - E[y|\boldsymbol{x}, \boldsymbol{z}, X = 0] \\
&= \text{Prob}(y = 1, r = 1|\boldsymbol{x}, \boldsymbol{z}) + \text{Prob}(y = 1, r = 0|\boldsymbol{x}, \boldsymbol{z}) \\
&- \text{Prob}(y = 0, r = 1|\boldsymbol{x}, \boldsymbol{z}) - \text{Prob}(y = 0, r = 0|\boldsymbol{x}, \boldsymbol{z}) \\
&= \boldsymbol{\Phi}_2(\boldsymbol{\beta}_2{}'\boldsymbol{x} + \gamma r, \boldsymbol{\beta}_1{}'\boldsymbol{z}, \rho|X = 1) + \boldsymbol{\Phi}_2(\boldsymbol{\beta}_2{}'\boldsymbol{x} + \gamma r, -\boldsymbol{\beta}_1{}'\boldsymbol{z}, -\rho|X = 1) \\
&- \boldsymbol{\Phi}_2[-(\boldsymbol{\beta}_2{}'\boldsymbol{x} + \gamma r), \boldsymbol{\beta}_1{}'\boldsymbol{z}, -\rho|X = 0] \\
&- \boldsymbol{\Phi}_2[-(\boldsymbol{\beta}_2{}'\boldsymbol{x} + \gamma r), -\boldsymbol{\beta}_1{}'\boldsymbol{z}, \rho|X = 0]
\end{aligned} \tag{9}$$

It is noted that the effect for the variable that appears only in $\boldsymbol{x}$ is interpreted as having

[9] Despite extensive testing in the model, the heterogeneity in the variance was not statistically significant.

[10] It is easier to interpret the marginal effects (compared to elasticities) for indicator variables (***Washington et al., 2020***).

a direct effect while the variable that appear only in $\boldsymbol{z}$ is interpreted as having an indirect effect on y that is transmitted through r. For the variable that appear in both $\boldsymbol{x}$ and $\boldsymbol{z}$, the total effect is combined and does not need to be treated as direct or indirect (***Greene, 2018***).

4.4 Results and discussion

In this study, separate probit models, the recursive bivariate probit model, and the random parameters recursive bivariate probit model with heterogeneity in the means are estimated. **Table 4-3** presents the results of the parameter estimation. The results indicate that the random parameters recursive bivariate probit model with heterogeneity in the means is superior in terms of Akaike information criterion (AIC). A detailed discussion of the results of parameter estimation among the models is provided in the sections that follow.

4.4.1 Effect of endogeneity

In this study, it was hypothesized that when modeling the relationship between crash type and injury severity, the endogeneity effect of crash type can be captured through the correlation coefficient in the recursive simultaneous model. As shown in **Table 4-3**, the significant correlation coefficients in all bivariate models provide statistical evidence for the correlation between the two structural disturbances. In other words, crash type may also be influenced by other factors that could affect the injury severity. The endogenous effect of crash type on injury severity is prevalent. However, the relationship between crash type and injury severity is complex and multifaceted. A possible explanation for the endogeneity of crash type is that unobserved factors including driver behavior and vehicle performance may affect both the crash type and injury outcome (***Mannering and Bhat, 2014***; ***Mannering et al., 2020***).

On the other hand, differences in the estimated parameters among independent probit models, recursive bivariate probit model and random parameter recursive bivariate probit model with heterogeneity in the means also justify the existence of endogeneity. For example, the likelihood of injury for a rollover crash would be underestimated if endogeneity were not considered (parameter estimate of 1.57 in the random parameters recursive bivariate model and 1.45 in the recursive bivariate model, compared with 0.85 in the independent model).

4.4.2 Influencing factors of crash type

Table 4-3 presents the results of parameter estimation for crash type (rollover crash). The likelihood of rollover crash is lower for cars. This is consistent with the finding of previous study that elevation of vehicles' center of gravity is positively associated with the probability of overturn (***Alrejjal et al., 2021***). Regarding the effects of environmental conditions, likelihood of rollover crashes is higher under dry road surface conditions (relative to wet and other road surface conditions) and daylight conditions (relative to dusk or dawn, and dark conditions). This may be because under favorable road and lighting conditions, drivers may drive faster and smugly (***Eluru and Bhat, 2007***). Regarding speed, studies have indicated that the speed of a vehicle plays a significant role in rollover intensity, as its kinetic energy affects the potential for vehicle rollover. Higher speeds are associated with an increased likelihood of rollover crashes (***Azimi et al., 2020***). Also, there is a tendency for drivers to adapt their speeds based on the posted speed limits, with higher speed limits often leading to increased driving speeds, a phenomenon known as speed generalization (***Elvik, 2015***). Hence, it is not surprising to observe that the likelihood of rollover crashes is higher for the ramp areas with higher speed limits (above 60 mph). Regarding the area type, the likelihood of rollover crashes is higher for ramps in rural areas. This could be due to the prevalence of hazards like hillslope and pavement edge drop-off at rural locations (***Islam and Pande, 2020***). With regard to horizontal alignment, the effect of straight ramp is random with a negative mean for rollover crash. As shown in **Figure 4.1(a)**, the normal distribution with the mean of -0.22 and the standard deviation of 0.33 indicates that likelihood of rollover crash is lower for 74.8% of observations. The negative parameters of female drivers and old drivers for straight ramp (-0.44 and -0.85, respectively) indicate that these two variables reduce the mean effect of straight ramp on rollover crash. This implies that female and old drivers are less likely involved in rollover crashes at straight ramps. Furthermore, from a vehicle dynamics standpoint, the likelihood of rollover crashes is smaller at straight ramps compared to curved ramps. This is due to the former's clear and predictable path, good visibility, and roadside hazard anticipation. Rollover crashes are often attributed to loss of vehicle control, reduced stability and over steering in emergency (***Martensen and Dupont, 2013***; ***Alrejjal et al., 2021***). Similarly, for interstate highways, the likelihood of rollover crashes is lower compared to non-interstate highways (US highway and State highway). Specifically, interstate highways are designed to accommodate higher volumes of traffic at relatively higher speeds, and generally provide efficient and safe transportation for long-distance travel and freight movement. Therefore, the higher standards for interstate highways, including lanes widths, access points frequency, and maintenance, contribute to a higher level of service compared to other highway classes (***AASHTO, 2010***, ***2018***). Thus, these characteristics may play a role in reducing the probability of rollover crashes at ramp areas of interstate road compared to those of non-interstates. Lastly, compared to road shoulders, the likelihood of rollover crashes is higher on traffic lanes. High-speed maneuvers of vehicles in traffic lanes, such as sudden lane changes, swerving, or avoiding obstacles, generally pose higher risk of rollover at those elements of the highway cross-section.

4.4.3 Influencing factors of injury severity

Table 4-3 presents the results of parameter estimation for injury severity. The following discussion focuses on the random parameters recursive bivariate probit model with heterogeneity in the means. **Table 4-4** presents the marginal effects of the exogenous variables for injury severity.

4.4.3.1 Driver characteristics

First, the probability of injury in the crash is higher when the driver is driving under the influence of alcohol and drugs. This is consistent with the findings of previous studies (***Wang et al., 2009***; ***Li et al., 2012***; ***Mergia et al., 2013***; ***Song et al., 2024b***). Second, dangerous driving behaviors including oversteering and aggressive driving behavior significantly increase the probability of injury, with the direct marginal effects of 0.0765 and 0.1533, respectively. Previous studies also indicated that aberrant driving behavior could increase the likelihood of more severe injury (***Paleti et al., 2010***; ***Song et al., 2024b***). To this end, he study result could help justify the need for driver education, training, and in-vehicle assistance system in enhancing the driver awareness, improving the defensive driving skills, and mitigating the crash injury risk at ramps (***Mallia et al., 2015***). Lastly, non-use of seat belt produces a random parameter with positive mean for injury severity. As shown in **Figure 4.1(b)**, the random parameter is normally distributed with the mean of 1.12 and standard deviation of 1.45. This implies that the likelihood of injury is higher for the non-use of seat belts for the majority of crashes (78.0%). In contrast, the negative parameter of female drivers (-0.50) indicates that the mean effect of non-use of seat belt on injury severity would decrease for female. While it may seem counterintuitive, this justifies the collective impact on the injury severity for the interference among factors including seat belt use and personal characteristics. For example, the level of emergency medical service (both quantity and quality) could mitigate the crash outcome. It is worth exploring the effects of response time and clearance time on the crash outcome when comprehensive traffic, crash, and trauma datasets are available (***Wong et al., 2007***; ***Tsui et al., 2009***; ***Castro et al., 2013***; ***Peura et al., 2015***).

4.4.3.2 Environmental conditions

As aforementioned, the results suggest that ramp area crashes under dry road surface and daylight conditions are more likely to be rollovers compared to wet and nighttime (including dusk and dawn) conditions. Hence, there is an indirect effect of road surface and lighting conditions on the injury severity mediated through crash type. As shown in **Table 4-4**, the probability of injury increases by 0.0256 for the crashes in dry road surface conditions and 0.0123 for crashes in daylight conditions. While favorable driving environment may provide a driver the better sense of safety and confidence, higher level of perceived safety would often result in complacency and more risky driving behaviors. There are compensatory effects of favorable road environment on the risk of rollover crashes and more severe injuries (***Kim et al., 2010***; ***Fountas et al., 2020***; ***Shaon and Qin, 2020***). Lastly, the effect of rural area is random with a negative mean. As shown in **Figure 4.1(c)**, the random parameter is normally distributed with the mean of -0.55 and standard deviation of 0.91. This implies that 72.7% of observations have a lower likelihood of injury. Despite the higher rollover crash likelihood for the crashes at rural road ramp areas, the probability of injury is reduced with the total marginal effect of -0.0411 could be explained by several reasons. First, rural areas typically have wider clear zones[11] in the US, offering more recovery space and buffer zones for vehicles in rollover incidents, potentially lessening injury severity. In addition, studies have demonstrated that the implementation of safety measures and infrastructure improvements on rural roads could play a role in reducing the likelihood of injury in crashes. For example, installing centerline rumble strips on rural roads can significantly decrease injury crashes (***Persaud et al., 2004***). These safety enhancements have the potential to mitigate crash severity and decrease the chances of sustaining injuries.

4.4.3.3 Roadway design

As the exogenous variables of horizontal alignment and speed limit are not significant in the equation of injury severity component, there exist indirect effects on the injury severity of crashes through the recursive structure. For example, the probability of injury for crashes at the straight ramp areas decreases by 0.0126, compared to that at

[11] A clear zone is an unobstructed and traversable roadside area for errant vehicles in the US (***AASHTO, 2011***).

the curved sections. In contrast, the probability of injury for crashes at the ramp areas with higher speed limit (above 60 mph) increases by 0.0097, compared to that with lower speed limits. Therefore, lowering the vehicle speed can reduce the probability of rollover crash and further reduce the possibility of injury due to crash. It is worth exploring the effectiveness of remedial measures (e.g., warning signs and variable speed limit) and Advanced Driver Assistance Systems (e.g., lane departure warning, curve warning and collision warning systems) in mitigating the risk of rollover crashes and enhancing safety of ramp areas (***Li et al., 2014***; ***Harper et al., 2016***; ***Wu et al., 2016***; ***Fleming et al., 2019***). Furthermore, the probability of injury for crashes on interstate highways decreases, with the total effect of 0.0442, compared to other road types. This is consistent with the findings of previous studies (***Chen et al., 2016***; ***Song et al., 2024b***).

4.4.3.4 Crash circumstances

Crash location on traffic lanes (compared to the road shoulder) has an opposite effect on the probabilities of rollover crashes and injury severities: the location on traffic lanes increases the likelihood of rollover crash by 0.42 but reduces that of injury by 0.12. Overall, as shown in **Table 4-4**, the total marginal effect of crash location is -0.0232. In addition, the effect of on-ramp on the probability of injury is random (normally distributed with the mean of -0.59 and standard deviation of 1.09), as shown in **Figure 4.1(d)**. This implies that there is a negative association between on-ramp and injury severity for the majority of crashes (70.6%). In contrast, the positive parameter of female driver (0.33) indicates that the mean effect of on-ramp on injury severity would increase for female. This justifies the heterogeneity in driver behavior among different genders at the highway ramp areas and their impacts on the association between crash location and injury severity (***Mannering et al., 2016***).

Table 4-3 Results of parameter estimation

Variable	Separate probit models		Recursive bivariate probit model		Random parameters recursive bivariate probit model with heterogeneity in means	
	Rollover	Injury	Rollover	Injury	Rollover	Injury
	Coefficient (*t*-stat)	Coefficient (*t*-stat)	Coefficient (*t*-stat)	Coefficient (*t*-stat)	Coefficient (*t*-stat)	Coefficient (*t*-stat)
Constant	-2.48 (-18.8)	-0.60 (-9.13)	-2.00 (-12.85)	-0.62 (-9.34)	-2.37 (-13.25)	-0.56 (-8.22)
Driver gender						
Female driver		0.16 (3.07)		0.17 (3.32)		
Alcohol or drugs						
Driving under the influence of alcohol or drugs		0.39 (4.52)		0.39 (4.53)		0.50 (5.10)
Aberrant driving behavior						
Oversteer		0.24 (2.94)		0.23 (2.84)		0.29 (3.23)
Aggressive driving		0.46 (4.52)		0.45 (4.35)		0.62 (5.16)
Safety belt						
Not used		0.72 (7.53)		0.70 (7.38)		1.12 (7.76)
Standard deviation						1.45 (9.68)
Vehicle type						
Car			-0.86 (-10.06)		-4.70 (-7.84)	
Standard deviation					2.78 (8.69)	
Road surface condition						

Variable	Separate probit models		Recursive bivariate probit model		Random parameters recursive bivariate probit model with heterogeneity in means	
	Rollover	Injury	Rollover	Injury	Rollover	Injury
	Coefficient (*t*-stat)	Coefficient (*t*-stat)	Coefficient (*t*-stat)	Coefficient (*t*-stat)	Coefficient (*t*-stat)	Coefficient (*t*-stat)
Dry	0.63 (7.15)		0.63 (6.48)		0.88 (7.20)	
Lighting condition						
Daylight	0.34 (4.34)		0.28 (3.37)		0.39 (3.98)	
Area type						
Rural	0.37 (4.49)	-0.16 (-3.01)	0.34 (3.63)	-0.19 (-3.50)	0.37 (3.53)	-0.55 (-7.11)
Standard deviation						0.91 (15.84)
Horizontal alignment						
Straight	-0.25 (-3.37)		-0.28 (-3.43)		-0.22 (-2.05)	
Standard deviation					0.33 (4.54)	
Speed limit						
Above 60 mph	0.25 (2.96)		0.22 (2.29)		0.33 (2.99)	
Road classification						
Interstate highway	-0.17 (-2.11)	-0.12 (-2.32)	-0.16 (-1.83)	-0.11 (-2.19)	-0.20 (-1.89)	-0.15 (-2.71)
Crash location						
On traffic lanes	0.34 (4.15)	-0.10 (-1.94)	0.32 (3.42)	-0.13 (-2.34)	0.42 (3.98)	-0.12 (-2.05)
Ramp type						
On-ramp		-0.14 (-2.84)		-0.14 (-2.80)		-0.59 (-7.49)
Standard deviation						1.09 (17.01)

Variable	Separate probit models		Recursive bivariate probit model		Random parameters recursive bivariate probit model with heterogeneity in means	
	Rollover	Injury	Rollover	Injury	Rollover	Injury
	Coefficient (*t*-stat)	Coefficient (*t*-stat)	Coefficient (*t*-stat)	Coefficient (*t*-stat)	Coefficient (*t*-stat)	Coefficient (*t*-stat)
Crash type (endogenous variable)						
Rollover crash		0.85 (8.85)		1.45 (5.11)		1.57 (7.47)
Heterogeneity in the mean of random parameter						
Female driver for the mean of non-use of seat belt						-0.50 (-1.95)
Female driver for the mean of rural area						0.31 (3.02)
Female driver for the mean of on-ramp						0.33 (3.11)
Female driver for the mean of straight alignment					-0.44 (-2.63)	
Older driver (age 60 or above) for the mean of straight alignment					-0.85 (-2.47)	
Correlation ρ			-0.34 (-2.14)		-0.31 (-2.21)	
Model performance						

Variable	Separate probit models		Recursive bivariate probit model		Random parameters recursive bivariate probit model with heterogeneity in means	
	Rollover	Injury	Rollover	Injury	Rollover	Injury
	Coefficient (*t*-stat)	Coefficient (*t*-stat)	Coefficient (*t*-stat)	Coefficient (*t*-stat)	Coefficient (*t*-stat)	Coefficient (*t*-stat)
Number of observations	3170	3170	3170		3170	
Degree of freedom	8	11	21		30	
Log-likelihood at convergence (LL(β))	-669.3459	-1756.0455	-2362.0346		-2341.5626	
AIC	1354.7	3534.1	4766.1		4743.1	

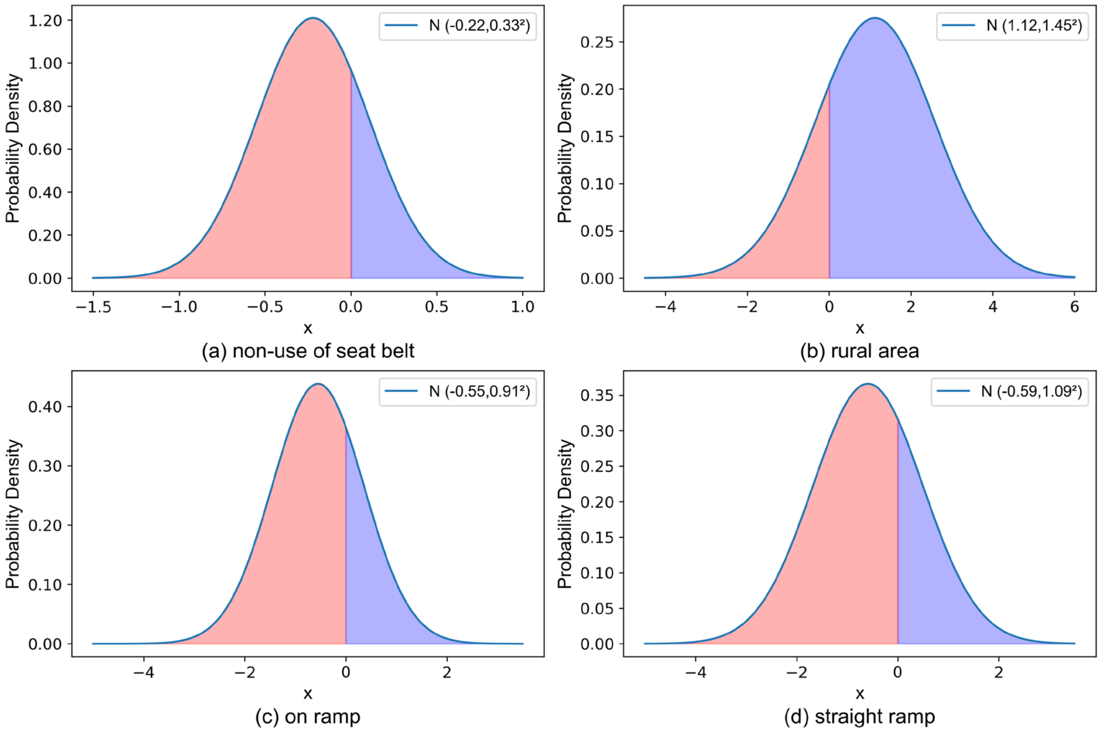


Figure 4.1 Distribution of random parameters

Table 4-4 Marginal effects for injury

Variable	**Indirect effect**	**Direct effect**	**Total effect**
Alcohol or drugs			
Driving under the influence of alcohol or drugs	-	0.1273	0.1273
Aberrant driving behavior			
Oversteer	-	0.0765	0.0765
Aggressive driving	-	0.1533	0.1533
Safety belt			
Not used	-	0.2530	0.2530
Vehicle type			
Car	-0.0413	-	-0.0413
Road surface condition			
Dry	0.0256	-	0.0256
Lighting condition			
Daylight	0.0123	-	0.0123
Area Type			
Rural	-	-	-0.0411
Horizontal alignment			

Variable	Indirect effect	Direct effect	Total effect
Straight	-0.0126	-	-0.0126
Speed limit			
Above 60 mph	0.0097	-	0.0097
Road classification			
Interstate highway	-	-	-0.0442
Crash location			
On traffic lanes	-	-	-0.0232
Ramp type			
On-ramp	-	-0.0442	-0.0442

4.5 Concluding remarks

Crash type is an important influencing factor that affects crash injury severity. In the conventional crash severity models, crash type is typically incorporated as an exogenous variable. However, an endogeneity effect (i.e., correlation between crash type and error term of the probability function of crash injury severity) is prevalent. To explore the endogeneity effect in a crash severity model, this study developed a random parameter recursive bivariate probit model with heterogeneity in the means for modeling crash injury severity of single-vehicle crashes at highway ramp areas. That way, the indirect effects of exogenous factors on injury severity through crash types can be accounted for. Furthermore, the effects of individual heterogeneity of the explanatory variables are considered in the simultaneous equation system.

The results indicate that the correlation of error terms in the simultaneous model is significant. This suggests the existence of endogenous effects of crash type on crash injury severity at ramp areas. The factors including driver characteristics, vehicle attributes, environmental conditions, roadway design, and crash circumstances that affect crash type and injury severity at ramp areas, are identified. For example, there exist significant effects for driving impairment and risky driving behavior, seat belt use, and road alignment on the likelihood of crash type and injury severity. These findings are expected to shed light on the development and implementation of effective remedial measures like driver training and education, variable speed limits, and warning signs to mitigate risk at hazardous ramp areas. Furthermore, the prevalence of endogeneity may be suggestive of the multifaceted nature of some road safety problems. Particularly, some exogenous variables that are significant only for the crash type are expected to have indirect effects on the crash severity (e.g., dry road surface conditions, daylight conditions) while some exogenous variables that are significant for both independent variables will exhibit opposite effects (e.g., rural area, crash on traffic lanes). This suggests that even where there exists endogenous effects for crash type on injury severity, certain interventions including guardrails and rumble strips installation could have direct effects on the probabilities of both rollover and injury crashes. Moreover, female drivers and old drivers are found to be statistically significant in the means of random parameters.

Chapter 5 Addressing Unobserved Heterogeneity at Road User Level for the Analysis of Conflict Risk at Toll Plaza

5.1 Introduction

Traffic and safety characteristics of toll plazas are different from that of other road entities because of the differences in geometric design, traffic management and control, and more importantly, weaving, diverging, and merging movements of traffic approaching the toll booths, especially vehicles slow down or stop to pay tolls when multiple toll collection methods (i.e., manual, and electronic) are available. Manual toll payment vehicles need to decelerate when approaching the toll booths, while electronic toll payment vehicles can continue to travel through the toll plaza at a relatively high speed. Derivation of the speed of mixed traffic can increase the crash risk of toll plaza (***Abdelwahab and Abdel-Aty, 2002***). It is crucial to identify the factors that affect the safety of toll plaza, hence effective countermeasures can be developed. Studies have assessed the safety of toll plaza, merging, and diverging areas, based on historical crash data. For example, toll plaza layout, horizontal curves, toll collection method, and traffic signs and road markings are found associated with the crash risk at toll plazas (***Wong et al., 2006***; ***Sze et al., 2008***; ***Abuzwidah et al., 2014***; ***Abuzwidah and Abdel-Aty, 2015***, ***2018***). In addition, lighting condition is associated with the occurrence and severity level of crash at diverging areas (***Mergia et al., 2013***). Furthermore, geometric design characteristics including number of lanes, road alignment, and length of deceleration lane are associated with the crash occurrence at off-ramp areas (***Chen et al., 2009***; ***Chen et al., 2011***; ***Calvi et al., 2012***). Last but not least, association between crash occurrence and possible risk factors can be moderated by collision type (i.e., rear-end, sideswipe, and angle collisions) (***Guo et al., 2019***).

In this study, modified traffic conflict indicator, taking into account vehicle length and width, angular and longitudinal movements, and conflict type (i.e., rear-end and sideswipe), is proposed to assess the safety risk at a tunnel toll plaza, based on high-resolution vehicle trajectory data obtained from drone video. Then, the correlated grouped random parameter multinomial logit approach with heterogeneity in the means of the random parameters is adopted to measure the association between conflict risk at tunnel toll plaza and possible factors, including vehicle class, speed and acceleration of vehicle, toll collection type, and spatial characteristics, for which effects of unobserved heterogeneity and correlation among random parameters at the road user level are accounted for.

The remainder of this paper is organized as follows. Data collection, model formulation, and analysis method are described in **Section 5.2**. Then, **Section 5.3** summarizes the data used. Furthermore, estimation results and interpretations are presented in **Section 5.4**. Finally, concluding remarks would be given in **Section 5.5**.

5.2 Method

5.2.1 Traffic conflict

Time-to-collision refers to the time required for two conflicting vehicles to collide if their speed and path remain unchanged (***Hayward, 1972***). For the rear-end collision of two vehicles travelling in the same direction, time-to-collision (TTC) can be calculated as,

$$TTC = \begin{cases} \dfrac{x_l - x_f - L_l}{v_f - v_l}, if\ v_f > v_l \\ \quad \infty, \qquad if\ v_f > v_l \end{cases} \tag{1}$$

where x_l is the displacement of front bumper of leading vehicle, x_f is that of following vehicle, v_l is the speed of leading vehicle, v_f is that of following vehicle, and L_l is the length of leading vehicle.

However, Equation (1) may not be capable of modeling the risk of angle and sideswipe collisions for diverging, merging, and weaving traffic. To this end, dimensions and angular movement of conflicting vehicles should be considered. **Figure 5.1** illustrates the typical interaction between two conflicting vehicles at the toll plaza, and diverging and merging areas. As shown in **Figure 5.1**, paths of Vehicle 1 and Vehicle 2 are intersecting at angle α. In addition, rectangle *1A1B1C1D* and *2A2B2C2D* denote the areas covered by Vehicle 1 and Vehicle 2, respectively. Shaped area covered by parallelogram *abcd* represents the overlapping area of trajectories of Vehicle 1 and 2 if their paths remain unchanged.

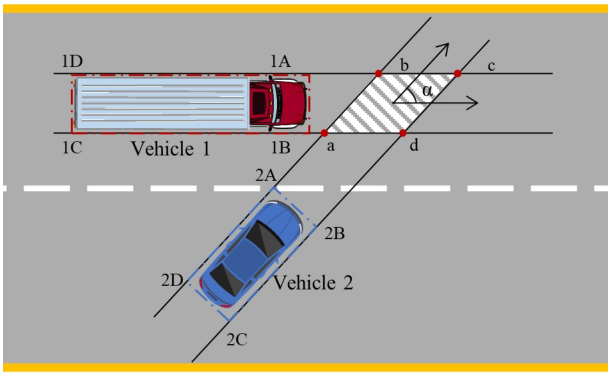


Figure 5.1 Illustration of interaction between two conflicting vehicles

Lengths, widths, and movement directions of Vehicles 1 and 2 can affect the location, shape, and size of the overlapping area. Point of contact for potential collision can be predicted, based on the assumption that Vehicles 1 and 2 would collide if their path and speed remain unchanged. Let t_{pq} denote the time at which the corner (p) of a vehicle reaches that (q) of the shaped area, where p is *1A*, *1B*, *1C*, …, and *2D*, and q is a, b, c, and d, respectively. There may be a collision when the front of one vehicle reaches the overlapping area before another vehicle completely leaves the area. For example, when $t_{1Ba} < t_{2Aa}$ and $t_{1Ca} > t_{2Aa}$, the front left corner (*2A*) of Vehicle 2 will hit the right side (*1B1C*) of Vehicle 1 at a (Interested reader is referred to **Figure A1** in the **Appendix** for all possible collision scenarios). Hence, time-to-collision can be estimated based on the difference in arrival time at the overlapping area between the conflicting vehicles, with which the two-dimensional vehicle motion is considered. One should note that a leading vehicle refers to the vehicle that arrives at the overlapping area first, based on the instantaneous motion of conflicting vehicles at the time of observation, in the subsequent analysis. Time step interval for the analysis depends on the frame rate of video (***Laureshyn et al., 2010***; ***Gu et al., 2019***).

In general, collision can be classified into four categories: (i) head-on collision, (ii) angle collision, (iii) sideswipe collision, and (iv) rear-end collision, based on the point of contact and intersecting angle of conflicting vehicles (***Wu et al., 2020***). In this study, maximum intersecting angle of the sample is less than nine degrees. Hence, only the sideswipe and rear-end conflicts are considered.

Figure 5.2 illustrates some possible conflict scenarios in this study. For example, rear-end conflict refers to that when the front of a vehicle hit the rear of another vehicle, and sideswipe conflict refers to that when the corner of a vehicle hit the side of another vehicle.

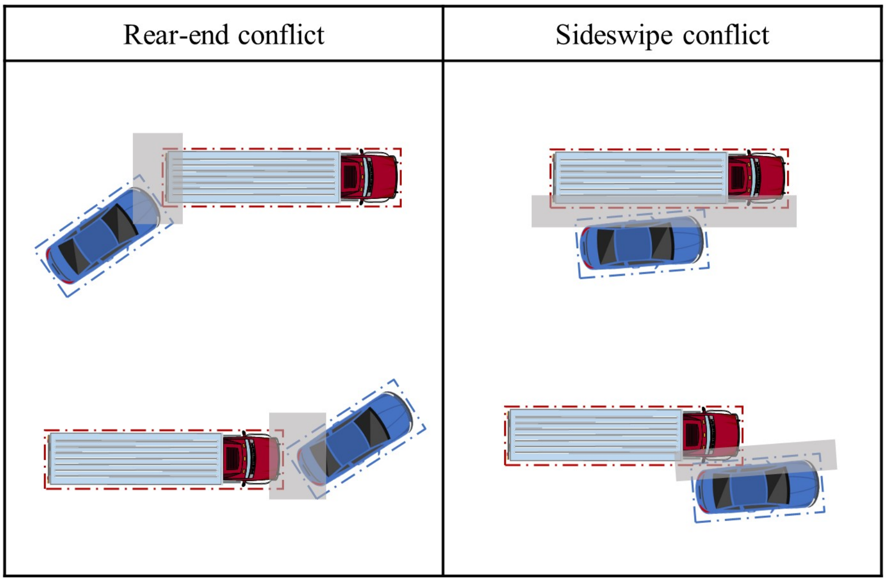


Figure 5.2 Illustration of possible conflict scenarios

In conventional traffic conflict analyses, minimum time-to-collision (***Johnsson et al., 2021***) and instantaneous time-to-collision (***Gu et al., 2019***; ***Xing et al., 2020a***) are commonly used to predict the occurrence of a traffic conflict. In this study, instantaneous time-to-collision is adopted as the conflict indicator, accounting for possible endogeneity of speed-related variables. When the value of instantaneous time-to-collision is lower than a pre-determined threshold, a traffic conflict will exist. In general, threshold of time-to-collision ranges from one to three seconds (***Madsen and Lahrmann, 2017***; ***Essa and Sayed, 2019***; ***Essa et al., 2019***; ***Wang et al., 2021***), and more than one stratification points can be established to distinguish among conflicts of different severity levels (***Essa et al., 2019***). In this study, two commonly used thresholds of time-to-collision are adopted: (i) 3 seconds for the occurrence of slight conflicts; and (ii) 1.5 seconds for the occurrence of severe conflicts. It should be noted that "conflict severity" refers to how close it is a collision may occur. It does not necessary imply the occurrence of severe crash. Therefore, it is more likely for a severe conflict to become a crash, compared to slight conflict (***Hydén, 1987***). Furthermore, a "traffic conflict" will be defined only when there is no traffic congestion (***Johnsson et al., 2021***) and the overlapping area is located within the pre-defined study area.

5.2.2 Model formulation

In this study, risk of traffic conflict at the toll plaza is modeled as discrete outcomes, namely (i) no conflict (i.e., normal interaction), (ii) slight conflict, and (iii) severe conflict. Hence, multinomial logit regression approach is adopted to measure the association between conflict risk and possible influencing factors, with "no conflict" defined as the baseline. For instance, probability of interaction i that has outcome j ($j \in J$) is given by:

$$P_{ij} = P\left(U_{ij} > U_{ik}\right), \qquad \forall k \neq j \tag{2}$$

where U_{ij} is the function that determines the probability of outcome j for interaction i

that is given by:

$$U_{ij} = \boldsymbol{\beta}_j{}'\boldsymbol{x}_i + \varepsilon_{ij}, i = 1, \ldots \ldots, n \tag{3}$$

where $\boldsymbol{x}_i$ is a vector of explanatory variables for interaction i, $\boldsymbol{\beta}_j$ is a vector of mean coefficients for outcome j, ε_{ij} is the error term which is assumed to be independent and identically distributed (IID) with Type 1 extreme value (Gumbel) distribution, and n is the total number of observations.

Then, the probability that outcome j will occur for interaction i can be expressed as:

$$\mathrm{P}_{ij}|\beta_j = \frac{\exp\left(\boldsymbol{\beta}_j{}'\boldsymbol{x}_i\right)}{\sum_{j=1}^{J} \exp\left(\boldsymbol{\beta}_j{}'\boldsymbol{x}_i\right)} \tag{4}$$

And, the unconditional probability can be computed as:

$$\mathrm{P}_{ij} = \int \frac{exp\left(\boldsymbol{\beta}_j{}'\boldsymbol{x}_i\right)}{\sum_{j=1}^{J} exp\left(\boldsymbol{\beta}_j{}'\boldsymbol{x}_i\right)} f(\boldsymbol{\beta}|\boldsymbol{\varphi})\, d\boldsymbol{\beta} \tag{5}$$

where $f(\boldsymbol{\beta}|\boldsymbol{\varphi})$ is the density function for vector $\boldsymbol{\beta}$, and $\boldsymbol{\varphi}$ is the vector of parameters that defines the density function.

In addition, the correlated random parameters approach with heterogeneity in the means is applied, accounting for the effect of unobserved heterogeneity. To account for the repeated observations from the same entity (i.e., vehicle interactions), grouped parameters approach is adopted. Parameters are allowed to vary across groups of observations. Hence, the coefficients would be modified as:

$$\boldsymbol{\beta}_{jr} = \boldsymbol{\beta}_j + \boldsymbol{\Theta}_j \boldsymbol{z}_j + \boldsymbol{\Gamma}\boldsymbol{\omega}_j \tag{6}$$

where $\boldsymbol{\beta}_{jr}$ is a vector of random parameters, $\boldsymbol{\Theta}_j$ is a matrix of estimated parameters, z_j is a vector of explanatory variables that capture heterogeneity in the means, $\boldsymbol{\omega}_j$ is normally distributed with $N(0, \sigma_j{}^2)$, and $\boldsymbol{\Gamma}$ is the Cholesky matrix based on Cholesky decomposition.

$\mathbf{\Gamma}$ matrix is a lower triangular matrix given as:

$$\mathbf{\Gamma} = \begin{bmatrix} \gamma_{1,1} & & & & \\ \gamma_{2,1} & \gamma_{2,2} & & & \\ \vdots & \vdots & \ddots & & \\ \gamma_{k-1,1} & \gamma_{k-1,2} & \cdots & \gamma_{k-1,k-1} & \\ \gamma_{k,1} & \gamma_{k,2} & \cdots & \gamma_{k,k-1} & \gamma_{k,k} \end{bmatrix} \tag{7}$$

The off-diagonal elements of $\mathbf{\Gamma}$ are usually set to be zero in conventional random parameter models. Thus, no correlation among random parameters is implied. To capture the possible correlations among random parameters, the off-diagonal elements should be non-zeros. One should note that the Cholesky matrix can be constrained (partially set to zero) (***Hensher et al., 2015***). The variance-covariance matrix can be written as:

$$\boldsymbol{C} = \mathbf{\Gamma}\mathbf{\Gamma}^T \tag{8}$$

which the diagonal elements are the standard deviations of random parameters, and off-diagonal elements are the covariance between random parameters, respectively. Standard deviation of the correlated random parameters can be expressed as:

$$\sigma_j = \sqrt{\gamma_{k,1}^2 + \gamma_{k,2}^2 + \cdots + \gamma_{k,k-1}^2 + \gamma_{k,k}^2} \tag{9}$$

t-statistics is used to assess the statistical significance of standard deviations of the correlated grouped random parameters. Standard error of the standard deviation is given by:

$$SE = \frac{S}{\sqrt{N}} \tag{10}$$

$$t = \frac{\sigma_j}{SE} \tag{11}$$

where S is the standard deviation, and N is the number of observations.

The correlation coefficient between two random parameters is computed as,

$$Cor(x_p, x_q) = \frac{cov(x_p, x_q)}{\sigma_p \sigma_q} \tag{12}$$

where $cov(x_p, x_q)$ is the covariance between the random parameters of variable x_p and x_q, and σ_p and σ_q are the standard deviations of the random parameters.

Likelihood-ratio test is used to assess the goodness-of-fit of two competing models with the chi-square test statistics given by,

$$X^2 = -2[LL(\beta_1) - LL(\beta_2)] \tag{13}$$

where $LL(\beta_1)$ and $LL(\beta_2)$ are the log-likelihood functions at convergence of Model 1 and Model 2, respectively, and degree of freedom is equal to the differences in number of parameters between competing models.

Parameters can be estimated by simulated maximum likelihood method. Previous studies indicate that 1,000 Halton draws is sufficient for the convergence of parameter estimation (***Meng et al., 2021***). Furthermore, marginal effects are also estimated to indicate the effects of explanatory variables on the outcome probabilities (***Washington et al., 2020***).

5.3 Data

Toll plaza (Kowloon bound) of Cross-Harbour Tunnel in Hong Kong (left hand driving rule applies) is selected as the study site. Cross-Harbour Tunnel, which was opened in 1972, is the busiest among the three underwater crossings connecting Kowloon and Hong Kong Island. In 2019, annual average daily traffic of Cross-Harbour Tunnel was 106,679 (***Hong Kong Transport Department, 2020***). As shown in **Figure 5.3**, number of lanes increase from three (near the tunnel portal) to eight (near the toll booths) when travelling along the toll plaza. Of the eight toll booths, three are allocated for electronic toll collection [i.e., Lane 1 (bus-only lane), Lane 2, and Lane 8], and five are for manual toll collection. Speed limit going through the electronic toll booths is 50 kph. As the driver behaviors and associated safety risks may change when drivers are approaching the toll booths from different distances, the study area is stratified into three (i.e., Zone 1, Zone 2, and Zone 3). In Zone 2, the number of lanes starts to increase, lane changing activities are frequent. In Zone 3, lane changing activities are partially restrained, particularly for electronic toll collection lanes.

To capture the aerial video, drone (DJI Mavic Air 2) is used in this study. Height of the drone is 100 meters above ground, and the field of view is 84 degrees. Observation survey was conducted during the daytime on 8 weekdays in October of 2020. Weather was fine (i.e., sunny and no wind) in the observation period. Also, there was no traffic jam. Overall, 120-minute video was captured. Resolution of the video was 1080p and frame rate was 30 fps. Vehicle trajectories were extracted from the video using the Automated Roadway Conflicts Identify System (ARCIS) of the University of Central Florida's Smart and Safe Transportation Lab (***Zheng et al., 2019***). For example, information on vehicle position (i.e., coordinates of centroid), dimensions (length and width), orientation, average speed, and acceleration rate can be obtained. For the calculation of TTC, readers may refer to the formulations given in **Figure A1** in the **Appendix**.

After the manual inspection and verification by experienced surveyors, trajectories of 2,217 vehicles were extracted. **Table 5-1** summaries the distribution of the sample, with respect to vehicle class and toll payment type. In particular, vehicles are classified into five categories: (i) private car, (ii) taxi, (iii) goods vehicle[12], (iv) bus, and (v) motorcycle. In addition, about half of the sample are using electronic toll payment (Count: 1,132; Proportion: 51.1%). **Figure 5.4** and **Figure 5.5** illustrate the

[12] In this study, goods vehicle refers to light goods vehicle (excluding van-type vehicle), medium goods vehicle, and heavy goods vehicle.

vehicle trajectories for different toll payment types and vehicle classes, respectively.

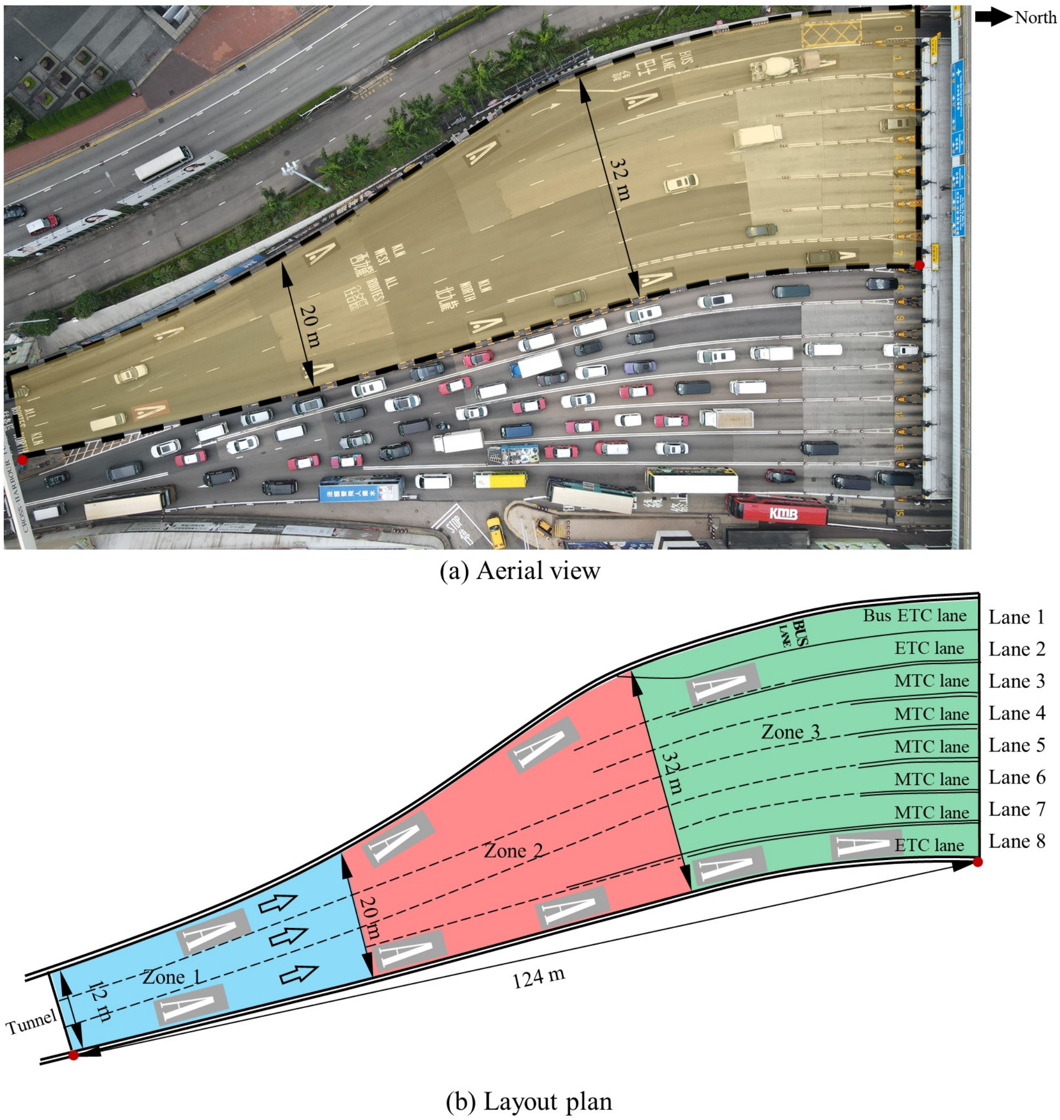


Figure 5.3 Layout of study site

Table 5-1 Distributions of the sample by toll payment type and vehicle class

Vehicle class	**Toll payment type**		**Overall**
	Manual	**Electronic**	
Private car	709 (32.0%)	835 (37.6%)	1544 (69.6%)
Taxi	175 (7.9%)	20 (0.9%)	195 (8.8%)
Goods vehicle	109 (4.9%)	101 (4.6%)	210 (9.5%)
Bus	5 (0.2%)	164 (7.4%)	169 (7.6%)
Motorcycle	87 (3.9%)	12 (0.5%)	99 (4.5%)
Total	**1085 (48.9%)**	**1132 (51.1%)**	**2217 (100.0%)**

It should be noted that the endogeneity issue is often overlooked in conflict analysis (***Yuan et al., 2022***). There has been trade-off between prediction and casusality for the inclusion of an endogenous variable in model estimation (***Mannering et al., 2020***). To capture as many explanatory variables that are recognized to affect safety (***Mannering and Bhat, 2014***), it is crucial to use speed-related variables in conflict analysis. Hence, optimal traffic management and control measures can be implemented to mitigate the real-time crash risk (***Formosa et al., 2020***; ***Mohammadian et al., 2021***; ***Fu and Sayed, 2022***). To address the endogeneity issue, while the instantaneous speed of leading and following vehicles are used to calculate time-to-collision, average speed in preceding one second (i.e., 30 frames) of conflict vehicles are used as explanatory variables in the model. **Table 5-2** shows the descriptive statistics of variables considered. Angular speed refers to the rate of change in vehicle direction (degree per second), where clockwise is considered as positive.

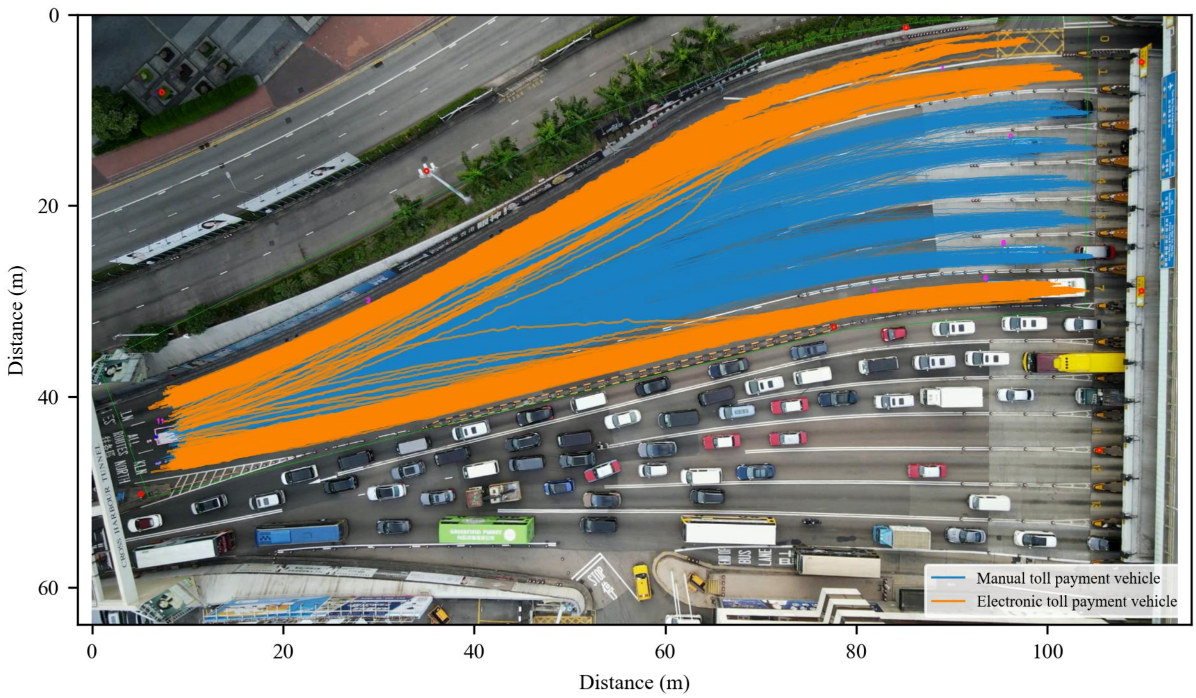
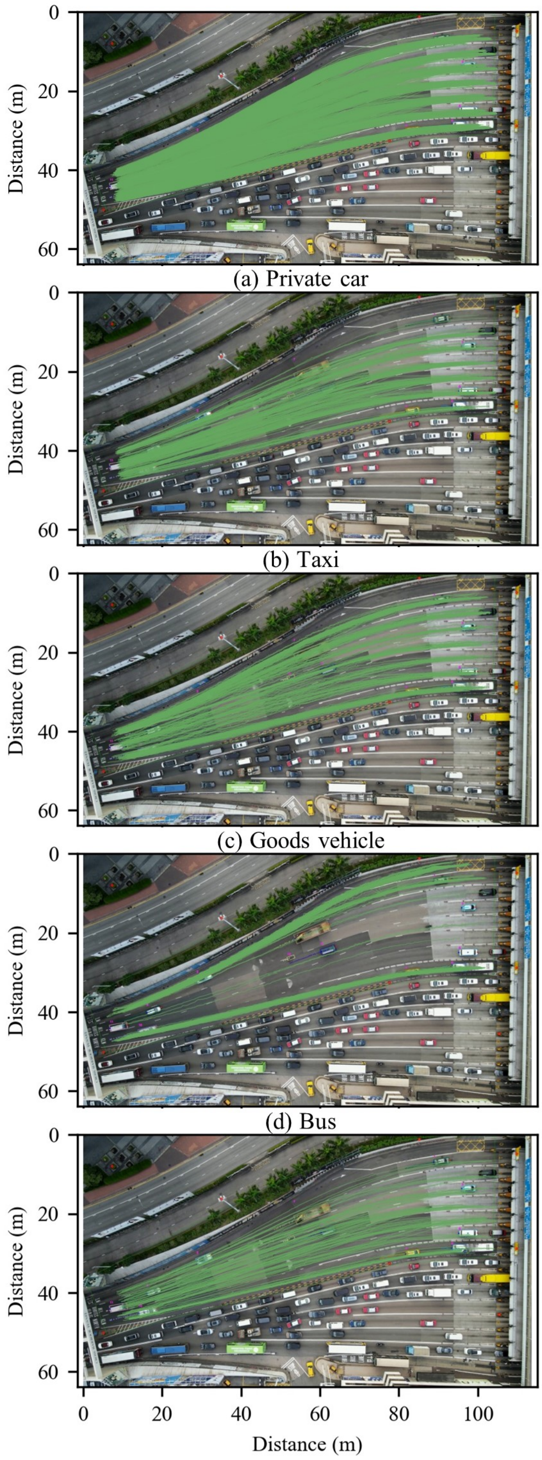


Figure 5.4 Vehicle trajectories for different toll payment types



(e) Motorcycle

Figure 5.5 Vehicle trajectories for different vehicle classes

Table 5-2 Descriptive statistics of variables considered

Factor		Rear-end interaction and conflict				Sideswipe interaction and conflict			
		Mean	S.D.	Min.	Max.	Mean	S.D.	Min.	Max.
At least one vehicle uses electronic toll payment = 1; Otherwise = 0		0.28	0.45	0	1	0.44	0.50	0	1
Zone 1		0.39	0.49	0	1	0.37	0.48	0	1
Zone 2		0.33	0.47	0	1	0.40	0.49	0	1
Zone 3		0.28	0.45	0	1	0.23	0.42	0	1
Leading vehicle	Average speed (meter/second)	8.56	3.17	0.59	21.6	11.01	2.79	2.46	19.54
	Acceleration (meter/second^2)	-0.82	1.04	-15.96	3.69	-0.44	2.01	-19.95	24.93
	Angular speed (degree/second)	4.46	2.55	0	34.18	4.75	3.07	0.0	28.36
	Private car (1 if yes, 0 otherwise)	0.66	0.47	0	1	0.68	0.47	0	1
	Taxi (1 if yes, 0 otherwise)	0.09	0.29	0	1	0.08	0.27	0	1
	Goods vehicle (1 if yes, 0 otherwise)	0.19	0.39	0	1	0.13	0.34	0	1
	Bus (1 if yes, 0 otherwise)	0.03	0.17	0	1	0.05	0.22	0	1
	Motorcycle (1 if yes, 0 otherwise)	0.03	0.17	0	1	0.06	0.23	0	1
Following vehicle	Average speed (meter/second)	12.41	2.60	2.53	20.86	11.81	2.75	1.41	24.47
	Acceleration (meter/second^2)	-0.26	3.10	-19.55	48.12	-0.24	3.88	-87.75	37.15
	Angular speed (degree/second)	4.32	2.96	0	28.82	5.04	3.61	0.0	28.27
	Private car (1 if yes, 0 otherwise)	0.63	0.48	0	1	0.66	0.48	0	1
	Taxi (1 if yes, 0 otherwise)	0.13	0.34	0	1	0.08	0.27	0	1
	Goods vehicle (1 if yes, 0 otherwise)	0.09	0.29	0	1	0.09	0.28	0	1
	Bus (1 if yes, 0 otherwise)	0.02	0.13	0	1	0.02	0.14	0	1
	Motorcycle (1 if yes, 0 otherwise)	0.13	0.33	0	1	0.16	0.36	0	1
Number of observations	No conflict	2726 (77.1%)				347 (24.5%)			
	Slight conflicts	698 (19.7%)				692 (48.8%)			
	Severe conflicts	111 (3.1%)				378 (26.7%)			
	Total observations	3535				1417			

5.4 Results and discussion

Correlated grouped random parameter models with heterogeneity in the means are established for rear-end and sideswipe conflicts respectively. **Table 5-3** shows the results of goodness-of-fit of correlated and uncorrelated grouped random parameter multinomial logit models with heterogeneity in the means for rear-end and sideswipe conflicts. Results of likelihood ratio test indicate that correlated grouped random parameter multinomial logit models (Chi-square test statistic is 21.58 for rear-end conflict and 10.76 for sideswipe conflict) are superior, compared to the uncorrelated models, both at the 0.01 level of significance. Results of parameter estimation, Cholesky matrix, and marginal effects are given in **Table 5-4** to **Table 5-10**.

Table 5-3 Model performance metrics between uncorrelated and correlated models

Metric	Rear-end conflicts model		Sideswipe conflicts model	
	Uncorrelated model	**Correlated model**	**Uncorrelated model**	**Correlated model**
McFadden R^2	0.591	0.594	0.253	0.257
Degrees of freedom	27	28	21	22
Log likelihood at convergence	-1588.24	-1577.45	-1162.69	-1157.31
AIC	3230.5	3210.9	2367.4	2358.6
Chi-square test statistic	21.58*		10.76*	

Note: * Statistical significance at the 0.01 level

5.4.1 Toll payment types

As shown in **Table 5-4**, likelihoods of rear-end conflicts (slight conflict: $\beta = 2.98$, severe conflict: $\beta = 1.89$) significantly increase when at least one electronic toll payment vehicle is involved. This could be because users of electronic toll payment tend to be more determined for the lane choice, and they may drive aggressively (they do not have to stop at the toll booths). As also shown in **Figure 5.3**, electronic toll collection lanes are located at the two sides of toll plaza (i.e., Lane 1, Lane 2, and Lane 8). Frequent lane changes may be required for the users of electronic toll payment (to reach the toll booths). On the other hand, users of manual toll payment may drive cautiously as they must keep checking the queue length of different toll booths, looking for that with the shortest queue (***Hong Kong Transport Department, 2020***). In addition, subscription fee is required for electronic toll payment in Hong Kong. Drivers using electronic toll payment are also frequent users of the tunnel. Hence, they should be more familiar with the route and traffic environment, compared to users of manual toll payment. Previous studies indicate that driver attention would decrease when route familiarity increases. Also, risk-taking behavior would be prevalent (***Martens and Fox, 2007***; ***Charlton and Starkey, 2013***; ***Weng et al., 2014***; ***Intini et al., 2018***).

5.4.2 Vehicle speed

For the effects of vehicle dynamics, likelihood of rear-end conflict (slight conflict: $\beta = -0.40$; severe conflict: $\beta = -0.67$) significantly decreases when the average speed of leading vehicle increases. In contrast, average speed of following vehicle is positively associated with the likelihood of rear-end conflict (slight conflict: $\beta = 0.22$; severe conflict: $\beta = 0.42$). This should align with the failure mechanism of traffic conflicts and crashes (***Xing et al., 2020a***; ***Zheng et al., 2021***). Just, it is important to recognize the potential endogeneity issue in real-time conflict analysis. In particular, one should be cautious about the model interpretation when speed-related variables, which were often used to calculate time-to-collision, are included as the explanatory variables in the model (***Mannering and Bhat, 2014***; ***Mannering et al., 2020***). In addition, effects of average speed on conflict risk are normally distributed. For example, as shown in **Table 5-5**, there is a 26.6% chance that likelihood of severe sideswipe conflict would increase when the average speed of leading vehicle increases. This is because vehicles involving in sideswipe conflicts are not necessarily travelling on the same traffic lane (***Jiménez et al., 2013***).

5.4.3 Vehicle acceleration rate

Furthermore, effects of acceleration rates on the likelihood of rear-end conflicts are captured. For instance, the likelihood of slight rear-end conflict is negatively associated with the acceleration rate of leading vehicle, while the risk of severe rear-end conflict would increase when the acceleration rate of following vehicle increases. This could be attributed to the geometric design and traffic characteristics of toll plaza area. As the toll plaza is close to the portal of underwater tunnel portal, and the road segment (a diverging area with frequent lane changing and weaving activities) connecting tunnel portal and toll booths is a moderate crest curve, sight distance can be constrained. Therefore, positive association between accelerate rate of following vehicle and conflict risk is prevalent. Such finding is indicative to the implementation of remedial measures like advanced warning signs, pavement markings, and rumble strips that can increase the awareness of driver and overall safety (***Wong et al., 2006***; ***Wong et al., 2012***). Effects of angular speeds on the likelihood of rear-end conflicts also show unobserved heterogeneity. For example, there is a 20.9% chance that the risk of severe rear-end conflict would increase when the angular speed of following vehicle increases. This could be attributed to the compensatory strategy adopted by the driver when one perceives that the lateral and steering controls are unstable (***Chen et al., 2021***), especially for the driver of following vehicle.

5.4.4 Vehicle class

For the effect of vehicle class, likelihood of rear-end conflict increases when the following vehicle is a goods vehicle. Also, likelihood of sideswipe conflict increases when the following vehicle is a good vehicle. This could be attributed to the prevalence of blind spots and reduced vision for the drivers of heavy goods vehicles. It is difficult for the drivers to observe surrounding traffic, especially for weaving and lane-changing. Therefore, awareness and attentiveness of drivers could be impaired (***Cook et al., 2011***; ***Marshall et al., 2020***). Such finding is indicative to vehicle design and innovations like advanced driver assistance system that could mitigate the risk attributed to reduced vision and inattentiveness (***Summerskill et al., 2016***). In addition, likelihood of sideswipe conflict of taxi is lower. This could be because of the compensatory behaviors adopted by the professional drivers who usually have better hazard perception skills (***Borowsky and Oron-Gilad, 2013***). Furthermore, risk of rear-end and sideswipe conflicts of motorcycle is higher. It could be because motorcyclists are more aggressive and risk-taking in general. As revealed in the crash statistics, crash involvement rates of motorcycle have been the highest among all vehicle classes in Hong Kong (***Hong Kong Transport Department, 2020***). Effective enforcement and educational strategies targeting to vulnerable driver groups, like commercial vehicle drivers and repeated conviction of traffic rules, can be implemented by the transport operators and government authorities. Therefore, safe driving can be promoted, and crash risk can be mitigated in the long run (***Chen et al., 2020***).

5.4.5 Spatial location

For the effect of vehicle's spatial location, likelihoods of rear-end and sideswipe conflicts at Zone 1 and Zone 2 are significantly lower, compared to Zone 3, at which lane changing activities are constrained. This may be contradictory to the findings of previous studies that crash risk is positively associated with lane changing and weaving activities at the merging areas (***Arbis and Dixit, 2019***; ***Gu et al., 2019***), diverging areas (***Xing et al., 2020b***), and road work zones (***Park et al., 2018***; ***Weng et al., 2018***). It could be because of the interactions between the vehicles stopped at Zone 3 to wait for the toll payment and those approaching from the tunnel portal. Speeds of the latter are usually higher. Also, driver capability to maintain lateral and steering stability could be impaired when the pavement markings are erased (***Chang et al., 2019***). Nevertheless, it is worth exploring on the association between geometric design, lane-changing activities, driver capability, and potential crash risk when information on visual perception and visual motor skills is available in naturalistic driving study and driving simulator experiment (***Chen et al., 2021***)

5.4.6 Unobserved heterogeneity

Furthermore, heterogeneity in the means of the random parameters is also considered. Factors that affect the means of the random parameters are identified. For example, for the likelihood of severe rear-end conflict, mean of the random parameter of angular speed is lower when an interaction is at Zone 1. For the likelihood of slight rear-end conflict, even that the mean coefficient of "following vehicle being a goods vehicle" is not significant ($\boldsymbol{\beta}_j$ being zero), heterogeneity in the mean $\boldsymbol{\Theta}_j$ is statistically significant ***(Alnawmasi and Mannering, 2019)***. Mean of the random parameter of "following vehicle being a goods vehicle" increases when an interaction is at Zone 1. For the likelihood of slight and severe sideswipe conflict, means of the random parameter of "average speed of following vehicle" increase when an interaction is at Zone 2. Heterogeneity in the mean estimation indicate the possible correlation between random parameter and exogeneous variables (***Mannering et al., 2016***). In this study, intervention effect by spatial location on the association between random parameter of vehicle dynamics and likelihood of conflict is revealed. Results are indicative to the real-time estimation of traffic conflict risk (***Alsaleh and Sayed, 2021***).

Correlations between random parameters are also considered. **Table 5-6** to **Table 5-9** present the estimates of Cholesky matrix and correlation coefficient matrix. For the rear-end conflicts, as shown in **Table 5-6**, there is no significant correlation between the random parameters for slight rear-end conflict ("average speed of leading vehicle" and "following vehicle being a goods vehicle"), off-diagonal elements for slight rear-end conflict are constrained to zero. As shown in **Table 5-7**, for the likelihood of severe rear-end conflict, random parameter of average speed of leading vehicle is negatively correlated to that of the angular speed of following vehicle ($\gamma = -0.672$, $Cor = -0.96$). This implies the diminishing effects for the unobserved heterogeneity of average speed of leading vehicle on that of the angular speed of following vehicle. For the sideswipe conflicts, as shown in **Table 5-8**, off-diagonal elements of Cholesky matrix are not constrained. As shown in **Table 5-9**, for the likelihood of severe sideswipe conflict, there is negative correlation between the random parameters of average speeds of leading and following vehicles ($\gamma = -0.305$, $Cor = -0.77$). Again, effects of unobserved heterogeneity of these two factors on severe sideswipe conflict are offsetting. Such findings may be attributed to the compensatory strategies adopted by the drivers under emergency, especially for experienced drivers (***Chen et al., 2021***). Yet, it is worth exploring the effectiveness of advanced driver assistance system in improving the driver performance and mitigating the safety risk, especially for lane changing and weaving activities in diverging area, when comprehensive data on driver visual perception, and perception-motor skills are available from naturalistic driving study and driving simulator experiment (***Chen et al., 2019b***; ***Chen et al., 2019a***; ***Mohammadian et al., 2021***).

Table 5-4 Results of parameter estimation of correlated model with heterogeneity in the means for rear-end conflicts

Variable	**Slight conflict**		**Severe conflict**	
	Coefficient	**t-stat**	**Coefficient**	**t-stat**
Constant	-1.73	-2.92	-7.19	-5.30
At least one electronic toll payment vehicle involved	2.98	6.73	1.89	2.19
Zone 1	-3.91	-10.22	-4.92	-4.39
Zone 2	-1.44	-6.07	-3.31	-4.57
Characteristics of leading vehicle				
Average speed	-0.40	-5.79	-0.94	-3.38
Standard deviation	0.45	11.45	0.78	4.37
Acceleration	-0.30	-3.82	-	-
Angular speed	-0.09	-2.76	-	-
Motorcycle	-	-	4.57	4.55
Characteristics of following vehicle				
Average speed	0.22	3.63	0.42	2.64
Acceleration	-	-	0.11	2.64
Angular speed	-	-	-0.57	-1.77
Standard deviation	-	-	0.70	2.78
Goods vehicle	0.18	0.41	2.11	2.22
Standard deviation	3.13	8.25	-	-
Motorcycle	1.74	5.50	3.19	4.84
Heterogeneity in the means of the random parameter				
Following vehicle being a goods vehicle: Zone 1	2.94	3.27	-	-
Angular speed of following vehicle: Zone 1	-	-	-0.59	-3.68
Model statistics				
McFadden R^2	0.594			
Number of observations	3535 (605 groups)			
Degree of freedom	28			
Log-likelihood at zero (LL(0))	-3883.59			

Variable	Slight conflict		Severe conflict	
	Coefficient	**t-stat**	**Coefficient**	**t-stat**
Log-likelihood at convergence (LL(β))	-1577.45			
Distributional effect of the random parameters across observations				
	Below zero	**Above zero**	**Below zero**	**Above zero**
Average speed of leading vehicle	81.3%	18.7%	88.6%	11.4%
Angular speed of following vehicle	-	-	79.1%	20.9%
Goods vehicle	47.7%	52.3%	-	-

Table 5-5 Results of parameter estimation of correlated model with heterogeneity in the means for sideswipe conflicts

Variable	Slight conflict		Severe conflict	
	Coefficient	**t-stat**	**Coefficient**	**t-stat**
Zone 1	-1.91	-5.15	-3.94	-6.74
Zone 2	-2.77	-2.82	-4.10	-2.62
Characteristics of leading vehicle				
Average speed	-	-	-0.25	-2.86
Standard deviation	-	-	0.40	4.42
Bus	-	-	3.23	3.78
Motorcycle	1.69	3.06	2.18	2.84
Characteristics of following vehicle				
Average speed	0.17	2.76	0.30	2.11
Standard deviation	0.13	8.53	0.40	4.95
Taxi	-	-	-1.92	-2.59
Goods vehicle	1.10	2.60	1.36	1.87
Motorcycle	0.90	1.93	-	-
Heterogeneity in the means of the random parameter				
Average speed of following vehicle: Zone 2	0.20	2.17	0.25	1.93
Model statistics				
McFadden R^2	0.257			
Number of observations	1417 (356 groups)			
Degree of freedom	22			
Log-likelihood at zero (LL(0))	-1556.73			
Log-likelihood at convergence (LL(β))	-1157.31			
Distributional effect of the random parameters across observations				
	Below zero	**Above zero**	**Below zero**	**Above zero**
Average speed of leading vehicle	-	-	73.4%	26.6%
Average speed of following vehicle	9.5%	90.5%	22.7%	77.3%

Table 5-6 Cholesky matrix of random parameters for rear-end conflict (t-statistic in parentheses)

Variable		Severe conflict: Average speed of leading vehicle	Severe conflict: Angular speed of following vehicle
Severe conflict	Average speed of leading vehicle	0.783 (4.37)	0
	Angular speed of following vehicle	-0.672 (-2.58)	0.207 (4.50)

Table 5-7 Correlation coefficient matrix of random parameters for rear-end conflict

Variable		Severe conflict: Average speed of leading vehicle	Severe conflict: Angular speed of following vehicle
Severe conflict	Average speed of leading vehicle	1.00	-0.96
	Angular speed of following vehicle	-0.96	1.00

Table 5-8 Cholesky matrix of random parameter for sideswipe conflict (t-statistic in parentheses)

Variable		Severe conflict: Average speed of leading vehicle	Severe conflict: Average speed of following vehicle
Severe conflict	Average speed of leading vehicle	0.397 (4.42)	0
	Average speed of following vehicle	-0.305 (-3.57)	0.254 (6.12)

Table 5-9 Correlation coefficient matrix of random parameters for sideswipe conflict

Variable		Severe conflict: Average speed of leading vehicle	Severe conflict: Average speed of following vehicle
Severe conflict	Average speed of leading vehicle	1.00	-0.77
	Average speed of following vehicle	-0.77	1.00

Table 5-10 Marginal effects on the probabilities of rear-end conflicts and sideswipe conflicts

Variable	Rear-end conflict			Sideswipe conflict		
	No conflict	Slight conflict	Severe conflict	No conflict	Slight conflict	Severe conflict
At least one electronic toll payment vehicle involved [SL]	-0.0548	**0.0578**	-0.0029	-	-	-
At least one electronic toll payment vehicle involved	-0.0042	-0.0019	**0.0060**	-	-	-

Variable	Rear-end conflict			Sideswipe conflict		
	No conflict	Slight conflict	Severe conflict	No conflict	Slight conflict	Severe conflict
[SE]						
Zone 1 [SL]	0.0673	**-0.0687**	0.0014	0.0930	**-0.1095**	0.0165
Zone 1 [SE]	0.0126	0.0018	**-0.0143**	0.0406	0.0340	**-0.0747**
Zone 2 [SL]	0.0323	**-0.0335**	0.0011	0.1020	**-0.1586**	0.0566
Zone 2 [SE]	0.0100	0.0026	-0.0127	0.0405	0.0836	**-0.1241**
Characteristics of leading vehicle						
Average speed [SL]	0.0649	**-0.0661**	0.0012	-	-	-
Average speed [SE]	-0.0144	-0.0042	**0.0186**	0.0438	0.0816	**-0.1255**
Acceleration [SL]	-0.0171	**0.0179**	-0.0008	-	-	-
Angular speed [SL]	0.0241	**-0.0252**	0.0011	-	-	-
Bus [SE]	-	**-**	-	-0.0047	-0.0110	**0.0157**
Motorcycle [SL]	-	**-**	-	-0.0066	**0.0120**	-0.0055
Motorcycle [SE]	-0.0021	-0.0005	**0.0026**	-0.0018	-0.0071	**0.0089**
Characteristics of following vehicle						
Average speed [SL]	-0.1719	**0.1795**	-0.0076	-0.2599	**0.4106**	-0.1507
Average speed [SE]	-0.0461	-0.0144	**0.0605**	-0.1412	-0.2733	**0.4145**
Acceleration [SE]	0.0001	0.0002	**-0.0003**	-	-	-
Angular speed [SE]	-0.0105	-0.0047	**0.0153**	-	-	-
Taxi [SE]	-	-	**-**	0.0029	0.0043	**-0.0072**
Goods vehicle [SL]	-0.0097	**0.0106**	-0.0009	-0.0087	**0.0145**	-0.0058
Goods vehicle [SE]	-0.0023	-0.0010	**0.0033**	-0.0026	-0.0072	**0.0098**
Motorcycle [SL]	-0.0156	**0.0168**	-0.0012	-0.0120	**0.0199**	-0.0079
Motorcycle [SE]	-0.0040	-0.0022	**0.0062**	-	-	-

Note: "SL" denotes slight conflict; "SE" denotes severe conflict; **bold values** indicate direct average marginal effects.

5.5 Concluding remarks

This study examines the safety risk of tunnel toll plaza based on the high-resolution trajectory data captured using drone. The correlated grouped random parameters multinomial logit model with heterogeneity in the means is adopted, accounting for the effects of repeated observations, unobserved heterogeneity, and correlation among random parameters at the road user level. Associations between possible influencing factors, occurrence and severity of traffic conflicts at the tunnel toll plaza are measured. In particular, modified traffic conflict indicator is proposed to account for the effects of dimensions (both width and length) and longitudinal and angular movement of interacting vehicles when estimating the conflict risk. This should improve the accuracy of conflict risk estimation, compared to the conventional (vehicle) centroid-based approach. In addition, effect of traffic conflict type (rear-end and sideswipe conflicts) on the association is considered.

Results indicate that when at least one electronic toll payment user is involved, likelihood of rear-end conflict would increase. As expected, likelihood of conflict is negatively associated with the average speed of leading vehicle, and positively associated with that of following vehicle. However, effects of average speed, acceleration rate, and angular speed on the conflict risks are random. These could be attributed to the compensatory behavior adopted by the drivers in emergency. Furthermore, conflict risks generally increase when goods vehicle and motorcycle are involved. This may be because of the reduced vision of goods vehicle drivers and risk-taking behaviors of motorcyclists. Nevertheless, correlated approach with heterogeneity in the means allows additional flexibility when capturing unobserved heterogeneity at the road user level. There are negative correlations between the random parameters of severe rear-end and sideswipe conflicts.

It is recommended that vehicle design could be enhanced, and advanced driver assistance system could be introduced to mitigate the risk attributed to the reduced vision of drivers, especially for heavy vehicles including buses and heavy goods vehicles. Findings are also indicative to the remedial design and measures for tunnel toll plazas including lane markings and advanced warning signs that can guide the drivers to the correct toll booths, and therefore reduce the risks of conflicts attributed to frequent lane changing and weaving activities.

Chapter 6 Conclusions and Recommendations

6.1 Conclusions

In this study, attempts have been made to assess the safety of highway merging and diverging areas using advanced econometric methods.

Chapter 2 has carried out a comprehensive literature review and gap analysis in the areas of highway safety analysis. The literature on the highway safety analysis at merging and diverging areas based on crash data is reviewed. Then, the surrogate safety measures and traffic conflict technique for safety analysis are discussed. Next, the literature with respect to analytic methods and critical methodological issues relating to highway safety analysis is reviewed. Finally, several research gaps are identified.

In Chapter 3, several factors that affect the crash injury severity of single-vehicle and multi-vehicle crashes at ramp areas are considered. Random parameters multinomial logit regression model with heterogeneity in means and variances is adopted to measure the association between possible influencing factors and crash severity at ramp areas based on the crash data from the North Carolina State of the United States in 2016-2018, with which the effects of unobserved heterogeneity and temporal instability are considered. Factors including driver characteristics, vehicle attributes, environmental conditions, roadway design, and crash circumstances are considered. Results indicate that there are considerable differences for the effects of aberrant driving, vehicle type, area type and crash location on the likelihood of injury between single-vehicle and multi-vehicle crashes. Additionally, there are opposite effects for the crashes in rural areas on the likelihood of injury between single-vehicle and multi-vehicle crashes. This justifies the need of developing and implementing targeted traffic control and management strategies that can reduce the risk of single-vehicle and multi-vehicle crashes separately. Furthermore, partially constrained modeling approach provides an efficient way to test for temporally shifting parameters by combining all data and defining parameters for each period. Results of partially constrained model and transferability assessments indicate that there is remarkable temporal instability. Effects of influencing factors on crash severity may change over time. By exploring the temporal transferability of crash injury severity model, understanding of the shifts in the effects of significant factors on the crash outcome could be enhanced.

In Chapter 4, the endogeneity effect in a crash severity model is explored. This study developed a random parameter recursive bivariate probit model with heterogeneity in the means and variances, for modeling crash injury severity of single-vehicle crashes at highway ramp areas. That way, the indirect effects of exogenous factors on injury severity through crash types can be accounted for. Furthermore, the effects of individual heterogeneity of the explanatory variables are considered in the simultaneous equation system. The results indicate that the correlation of error terms in the simultaneous model is significant. This justifies the endogenous effect of crash type on crash injury severity at ramp areas. The factors including driver characteristics, vehicle attributes, environmental conditions, roadway design, and crash circumstances that affect crash type and injury severity at ramp areas, are revealed. For example, there exist significant effects for driver gender, driving impairment and risky driving behavior, seat belt use, and road alignment on the likelihood of crash type and injury severity. This should shed light on the development and implementation of effective remedial measures like driver training and education, variable speed limits, and warning signs to mitigate the risk at hazardous locations at ramps. Furthermore, the prevalence of endogeneity may imply the multifaceted nature of some road safety problems. Even if there is endogenous effect for crash type on injury severity, some interventions including guardrails and rumble strips can have direct effects on the probabilities of both rollover and injury crashes.

In Chapter 5, the safety risk of tunnel toll plaza is examined based on the high-resolution trajectory data captured using drone. The correlated grouped random parameters multinomial logit model with heterogeneity in the means is adopted, accounting for the effects of repeated observations, unobserved heterogeneity, and correlation among random parameters at the road user level. Associations between possible influencing factors, occurrence and severity of traffic conflicts at the tunnel toll plaza are measured. In particular, modified traffic conflict indicator is proposed to account for the effects of dimensions (both width and length) and longitudinal and angular movement of interacting vehicles when estimating the conflict risk. This should improve the accuracy of conflict risk estimation, compared to the conventional (vehicle) centroid-based approach. In addition, effect of traffic conflict type (rear-end and sideswipe conflicts) on the association is considered. Results indicate that when at least one electronic toll payment user is involved, likelihood of rear-end conflict would increase. As expected, likelihood of conflict is negatively associated with the average speed of leading vehicle, and positively associated with that of following vehicle. However, effects of average speed, acceleration rate, and angular speed on the conflict risks are random. These could be attributed to the compensatory behavior adopted by the drivers in emergency. Furthermore, conflict risks generally increase when goods vehicle and motorcycle are involved. This may be because of the reduced vision of goods vehicle drivers and risk-taking behaviors of motorcyclists. Nevertheless, correlated approach with heterogeneity in the means allows additional flexibility when capturing unobserved heterogeneity at the road user level. There are negative correlations between the random parameters of severe rear-end and sideswipe conflicts. It is recommended that vehicle design could be enhanced, and advanced driver assistance system could be introduced to mitigate the risk attributed to the reduced vision of drivers, especially for heavy vehicles including buses and heavy goods vehicles. Findings are also indicative to the remedial design and measures for tunnel toll plazas including lane markings and advanced warning signs that can guide the drivers to the correct toll booths, and therefore reduce the risks of conflicts attributed to frequent lane changing and weaving activities.

6.2 Findings and contributions

This research has made significant contributions to the fields of safety analysis of highway merging and diverging areas through three studies.

The study on transferability of crash injury severity model at ramps contributes to the field by addressing the issues regarding the accuracy and reliability of crash injury severity models. The contribution of this study is twofold. Firstly, differences in the association measure of injury severity at ramp areas between single-vehicle and multi-vehicle crashes are evaluated. Results indicate that there are considerable differences for the effects of aberrant driving, vehicle type, area type and crash location on the likelihood of injury between single-vehicle and multi-vehicle crashes. Secondly, issues of unobserved heterogeneity and transferability for the analysis of crash injury severity at ramp areas are addressed. Results of partially constrained model and transferability assessments indicate that there is remarkable temporal instability. Effects of influencing factors on crash severity may change over time.

The study on the endogeneity effect in a crash severity model contributes to the field by examining the indirect effects of exogenous factors on injury severity through crash types. The results indicate that the correlation of error terms in the simultaneous model is significant. This justifies the endogenous effect of crash type on crash injury severity at ramp areas. The factors including driver characteristics, vehicle attributes, environmental conditions, roadway design, and crash circumstances that affect crash type and injury severity at ramp areas, are revealed. In addition, a random parameter recursive bivariate probit model with heterogeneity in the means and variances is developed to explore the effects of individual heterogeneity of the explanatory variables in the simultaneous equation system.

The study on unobserved heterogeneity at the road user level contributes to the field by exploring associations between possible influencing factors and severity of traffic conflicts at the tunnel toll plaza. In particular, modified traffic conflict indicator is proposed to account for the effects of dimensions (both width and length) and longitudinal and angular movement of interacting vehicles when estimating the conflict risk. In addition, the effect of traffic conflict type (rear-end and sideswipe conflicts) on the association is considered. Furthermore, correlated approach with heterogeneity in the means allows additional flexibility when capturing unobserved heterogeneity at the road user level.

6.3 Limitations and recommendations

6.3.1 Limitations

Despite the contributions to literature, the discussion of this research should take into account its limitations.

Regarding the study of crash injury severity model, information on real-time traffic characteristics (e.g., traffic flow and speed) and vehicle conditions (e.g., vehicle motion and position) that could affect ramp-area crash type and injury severity, are unavailable. They are crucial for the prediction of crash circumstances, impact force in the collision, and thus the crash outcome. In particular, spatial transferability allows for the application of models developed in one geographic location to be effectively used in other areas, enhancing the generalizability and applicability of the findings.

However, due to data limitations, the heterogeneity and stability of factors influencing injury severity in different spatial contexts is not examined in this study. Moreover, problems of some zero observations and imbalanced crash data (for more severe injury crashes) are prevalent.

With respect to the traffic conflict analysis, this study is limited to a single toll plaza area with specific geometry, lane, and toll booth configuration in Hong Kong. To generalize the findings for practical applications, it is worth investigating for the effects of geometric design, and configurations (i.e., lane allocation, traffic signs, pavement markings, and speed limit) on the risk of conflict when the trajectory data at other locations are available. Furthermore, information about driver characteristics obtained from the video is very limited. Effects of driver attention, visual perception, and perception-motor skill could be explored when comprehensive driving data are obtained from naturalistic driving study and driving simulator experiment in the future. Last but not least, a widespread concern in conflict-based analysis is that there may be endogeneity issues in statistical modeling as most of the conflict indicators are computed from vehicle speed and distances between vehicles. It is crucial to address the endogeneity issue in the model using appropriate statistical corrections (***Guevara and Ben-Akiva, 2012***; ***Song et al., 2024a***).

6.3.2 Recommendations for future research

The current work can be further extended in the future, contributing to a more comprehensive and nuanced understanding of highway safety at merging and diverging areas.

In future study, it is worth exploring the transferable real-time safety evaluation models for proactive road safety. The transferability of real-time evaluations allows for identification of potential safety hazards across different regions or time periods before accidents. For instance, developing transferability procedures for highway safety models can enable the adaptation of successful safety prediction models from one location to another, enhancing the efficiency of safety assessments on diverse road networks. In addition, by providing real-time safety status analysis results based on current traffic conditions, authorities can take timely and effective measures to address safety concerns and ensure the well-being of road users.

In addition, it may be worth exploring the effect of vehicle maneuvers including lane changing and weaving on the crash severity when comprehensive vehicle trajectory data are available. Furthermore, advanced data analytic approaches (e.g., generative models) can be adopted to throw more light towards the resolution of the problem of data imbalance.

References

AASHTO, 2010. Highway safety manual. American Association of State Highway and Transportation Officials, Washington, D.C.

AASHTO, 2011. Roadside design guide. American Association of State Highway and Transportation Officials, Washington, D.C.

AASHTO, 2018. A policy on geometric design of highways and streets. American Association of State Highway and Transportation Officials, Washington, D.C.

Abay, K.A., Paleti, R., Bhat, C.R., 2013. The joint analysis of injury severity of drivers in two-vehicle crashes accommodating seat belt use endogeneity. Transportation Research Part B 50, 74-89.

Abdel-Aty, M., 2003. Analysis of driver injury severity levels at multiple locations using ordered probit models. Journal of Safety Research 34 (5), 597-603.

Abdelwahab, H., Abdel-Aty, M., 2002. Artificial neural networks and logit models for traffic safety analysis of toll plazas. Transportation Research Record 1784, 115-125.

Abuzwidah, M., Abdel-Aty, M., Ahmed, M.M., 2014. Safety evaluation of hybrid main-line toll plazas. Transportation Research Record 2435, 53-60.

Abuzwidah, M., Abdel-Aty, M., 2015. Safety assessment of the conversion of toll plazas to all-electronic toll collection system. Accident Analysis and Prevention 80, 153-161.

Abuzwidah, M., Abdel-Aty, M., 2018. Crash risk analysis of different designs of toll plazas. Safety Science 107, 77-84.

Afghari, A.P., Washington, S., Haque, M.M., Li, Z., 2018. A comprehensive joint econometric model of motor vehicle crashes arising from multiple sources of risk. Analytic Methods in Accident Research 18, 1-14.

Afghari, A.P., Faghih Imani, A., Papadimitriou, E., van Gelder, P., Hezaveh, A.M., 2021. Disentangling the effects of unobserved factors on seatbelt use choices in multi-occupant vehicles. Journal of Choice Modelling 41, 100324.

Afghari, A.P., Papadimitriou, E., Pilkington-Cheney, F., Filtness, A., Brijs, T., Brijs, K., Cuenen, A., De Vos, B., Dirix, H., Ross, V., Wets, G., Lourenço, A., Rodrigues, L., 2022. Investigating the effects of sleepiness in truck drivers on their headway: An instrumental variable model with grouped random parameters and heterogeneity in their means. Analytic Methods in Accident Research 36, 100241.

Ahmed, S.S., Pantangi, S.S., Eker, U., Fountas, G., Still, S.E., Anastasopoulos, P.C., 2020. Analysis of safety benefits and security concerns from the use of autonomous vehicles: A grouped random parameters bivariate probit approach with heterogeneity in means. Analytic Methods in Accident Research 28, 100134.

Ahmed, S.S., Cohen, J., Anastasopoulos, P.C., 2021. A correlated random parameters with heterogeneity in means approach of deer-vehicle collisions and resulting injury-severities. Analytic Methods in Accident Research 30, 100160.

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. A temporal assessment of distracted driving injury severities using alternate unobserved-heterogeneity modeling approaches. Analytic Methods in Accident Research 34, 100216.

Alnawmasi, N., Mannering, F., 2023. An analysis of day and night bicyclist injury severities in vehicle/bicycle crashes: A comparison of unconstrained and partially constrained temporal modeling approaches. Analytic Methods in Accident Research 40, 100301.

Alogaili, A., Mannering, F., 2022. Differences between day and night pedestrian-injury severities: Accounting for temporal and unobserved effects in prediction. Analytic Methods in Accident Research 33, 100201.

Alrejjal, A., Farid, A., Ksaibati, K., 2021. A correlated random parameters approach to investigate large truck rollover crashes on mountainous interstates. Accident Analysis and Prevention 159, 106233.

Alsaleh, R., Sayed, T., 2021. Markov-game modeling of cyclist-pedestrian interactions in shared spaces: A multi-agent adversarial inverse reinforcement learning approach. Transportation Research Part C 128, 103191.

Anastasopoulos, P.C., Mannering, F., 2011. An empirical assessment of fixed and random parameter logit models using crash- and non-crash-specific injury data. Accident Analysis and Prevention 43 (3), 1140-1147.

Arbis, D., Dixit, V.V., 2019. Game theoretic model for lane changing: Incorporating conflict risks. Accident Analysis and Prevention 125, 158-164.

Arun, A., Haque, M.M., Washington, S., Sayed, T., Mannering, F., 2021a. A systematic review of traffic conflict-based safety measures with a focus on application context. Analytic Methods in Accident Research 32, 100185.

Arun, A., Haque, M.M., Bhaskar, A., Washington, S., Sayed, T., 2021b. A bivariate extreme value model for estimating crash frequency by severity using traffic conflicts. Analytic Methods in Accident Research 32, 100180.

Arun, A., Haque, M.M., Bhaskar, A., Washington, S., Sayed, T., 2021c. A systematic mapping review of surrogate safety assessment using traffic conflict techniques. Accident Analysis and Prevention 153, 106016.

Arun, A., Haque, M.M., Bhaskar, A., Washington, S., 2022. Transferability of multivariate extreme value models for safety assessment by applying artificial intelligence-based video analytics. Accident Analysis and Prevention 170, 106644.

Azimi, G., Rahimi, A., Asgari, H., Jin, X., 2020. Severity analysis for large truck rollover crashes using a random parameter ordered logit model. Accident Analysis and Prevention 135, 105355.

Behnood, A., Mannering, F., 2015. The temporal stability of factors affecting driver-injury severities in single-vehicle crashes: Some empirical evidence. Analytic Methods in Accident Research 8, 7-32.

Behnood, A., Mannering, F., 2017. The effect of passengers on driver-injury severities in single-vehicle crashes: A random parameters heterogeneity-in-means approach. Analytic Methods in Accident Research 14, 41-53.

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, 100102.

Bhat, C.R., Born, K., Sidharthan, R., Bhat, P.C., 2014. A count data model with endogenous covariates: Formulation and application to roadway crash frequency at intersections. Analytic Methods in Accident Research 1, 53-71.

Boggess, B.M., Morr, D.R., Peterman, E.K., Wiechel, J.F., 2010. Experimental evaluation of underride analysis techniques and empirical validation of a new analytical technique. Accident Analysis and Prevention 42 (1), 140-152.

Borowsky, A., Oron-Gilad, T., 2013. Exploring the effects of driving experience on hazard awareness and risk perception via real-time hazard identification, hazard classification, and rating tasks. Accident Analysis and Prevention 59, 548-565.

Calvi, A., Benedetto, A., De Blasiis, M.R., 2012. A driving simulator study of driver performance on deceleration lanes. Accident Analysis and Prevention 45, 195-203.

Castro, M., Paleti, R., Bhat, C.R., 2013. A spatial generalized ordered response model to examine highway crash injury severity. Accident Analysis and Prevention 52, 188-203.

Chang, F., Yasmin, S., Huang, H., Chan, A.H.S., Haque, M.M., 2022. Modeling endogeneity between motorcyclist injury severity and at-fault status by applying a Bayesian simultaneous random-parameters model with a recursive structure. Analytic Methods in Accident Research 36, 100245.

Chang, K., Ramirez, M.V., Dyre, B., Mohamed, M., Abdel-Rahim, A., 2019. Effects of longitudinal pavement edgeline condition on driver lane deviation. Accident Analysis and Prevention 128, 87-93.

Chang, L.-Y., Chien, J.-T., 2013. Analysis of driver injury severity in truck-involved accidents using a non-parametric classification tree model. Safety Science 51 (1), 17-22.

Charlton, S.G., Starkey, N.J., 2013. Driving on familiar roads: Automaticity and inattention blindness. Transportation Research Part F 19, 121-133.

Chen, C., Zhang, G., Qian, Z., Tarefder, R.A., Tian, Z., 2016. Investigating driver injury severity patterns in rollover crashes using support vector machine models. Accident Analysis and Prevention 90, 128-139.

Chen, H., Liu, P., Lu, J.J., Behzadi, B., 2009. Evaluating the safety impacts of the number and arrangement of lanes on freeway exit ramps. Accident Analysis and Prevention 41 (3), 543-51.

Chen, H., Zhou, H., Zhao, J., Hsu, P., 2011. Safety performance evaluation of left-side off-ramps at freeway diverge areas. Accident Analysis and Prevention 43 (3), 605-12.

Chen, K., Xu, C., Liu, P., Li, Z., Wang, Y., 2024. Evaluating the performance of traffic conflict measures in real-time crash risk prediction using pre-crash vehicle trajectories. Accident Analysis and Prevention 203, 107640.

Chen, S., Saeed, T.U., 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, T., Sze, N.N., Bai, L., 2019a. Safety of professional drivers in an ageing society – A driving simulator study. Transportation Research Part F 67, 101-112.

Chen, T., Bai, L., Sze, N.N., 2019b. Factors affecting the severity of rear-end conflicts: a driving simulator study. 2019 5th International Conference on Transportation Information and Safety (ICTIS), 1182-1187.

Chen, T., Sze, N.N., Saxena, S., Pinjari, A.R., Bhat, C.R., Bai, L., 2020. Evaluation of penalty and enforcement strategies to combat speeding offences among professional drivers: A Hong Kong stated preference experiment. Accident Analysis and Prevention 135, 105366.

Chen, T., Sze, N.N., Newnam, S., Bai, L., 2021. Effectiveness of the compensatory strategy adopted by older drivers: Difference between professional and non-professional drivers. Transportation Research Part F 77, 168-180.

Cook, S., Summerskill, S., Marshall, R., Richardson, J.H., Lawton, C., Grant, R., Bayer, S.H., Lenard, J., Clemo, K., 2011. The development of improvements to drivers' direct and indirect vision from vehicles - Phase 2. Loughborough University. ***https://hdl.handle.net/2134/8873***.

Dzinyela, R., Alnawmasi, N., Kofi Adanu, E., Dadashova, B., Lord, D., Mannering, F., 2024. A multi-year statistical analysis of driver injury severities in single-vehicle freeway crashes with and without airbags deployed. Analytic Methods in Accident Research 41, 100317.

El-Basyouny, K., Sayed, T., 2013. Safety performance functions using traffic conflicts. Safety Science 51 (1), 160-164.

Eluru, N., Bhat, C.R., 2007. A joint econometric analysis of seat belt use and crash-related injury severity. Accident Analysis and Prevention 39 (5), 1037-1049.

Elvik, R., 2015. A statistical law in the perception of risks and physical quantities in traffic. Accident Analysis and Prevention 82, 36-44.

Essa, M., Sayed, T., Reyad, P., 2019. Transferability of real-time safety performance functions for signalized intersections. Accident Analysis and Prevention 129, 263-276.

Essa, M., Sayed, T., 2019. Full Bayesian conflict-based models for real time safety evaluation of signalized intersections. Accident Analysis and Prevention 129, 367-381.

Feng, S., Li, Z., Ci, Y., Zhang, G., 2016. Risk factors affecting fatal bus accident severity: Their impact on different types of bus drivers. Accident Analysis and Prevention 86, 29-39.

Filippini, M., Greene, W.H., Kumar, N., Martinez-Cruz, A.L., 2018. A note on the different interpretation of the correlation parameters in the Bivariate Probit and the Recursive Bivariate Probit. Economics Letters 167, 104-107.

Fleming, J.M., Allison, C.K., Yan, X., Lot, R., Stanton, N.A., 2019. Adaptive driver modelling in ADAS to improve user acceptance: A study using naturalistic data. Safety Science 119, 76-83.

Formosa, N., Quddus, M., Ison, S., Abdel-Aty, M., Yuan, J., 2020. Predicting real-time traffic conflicts using deep learning. Accident Analysis and Prevention 136, 105429.

Fountas, G., Anastasopoulos, P.C., 2017. A random thresholds random parameters hierarchical ordered probit analysis of highway accident injury-severities. Analytic Methods in Accident Research 15, 1-16.

Fountas, G., Sarwar, M.T., Anastasopoulos, P.C., Blatt, A., Majka, K., 2018a. Analysis of stationary and dynamic factors affecting highway accident occurrence: A dynamic correlated grouped random parameters binary logit approach. Accident Analysis and Prevention 113, 330-340.

Fountas, G., Anastasopoulos, P.C., Abdel-Aty, M., 2018b. Analysis of accident injury-severities using a correlated random parameters ordered probit approach with time variant covariates. Analytic Methods in Accident Research 18, 57-68.

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.

Fountas, G., Fonzone, A., Olowosegun, A., McTigue, C., 2021. Addressing unobserved heterogeneity in the analysis of bicycle crash injuries in Scotland: A correlated random parameters ordered probit approach with heterogeneity in means. Analytic Methods in Accident Research 32, 100181.

Fu, C., Sayed, T., 2022. Bayesian dynamic extreme value modeling for conflict-based real-time safety analysis. Analytic Methods in Accident Research 34, 100204.

Geedipally, S.R., Lord, D., 2010. Investigating the effect of modeling single-vehicle and multi-vehicle crashes separately on confidence intervals of Poisson–gamma models. Accident Analysis and Prevention 42 (4), 1273-1282.

Geedipally, S.R., Bonneson, J.A., Pratt, M.P., Lord, D., 2014. Analysis of injury severity in crashes on ramps and at crossroad ramp terminals. Transportation Research Record 2435, 37-44.

Gong, H., Fu, T., Sun, Y., Guo, Z., Cong, L., Hu, W., Ling, Z., 2022. Two-vehicle driver-injury severity: A multivariate random parameters logit approach. Analytic Methods in Accident Research 33, 100190.

Greene, W.H., 1998. Gender economics courses in liberal arts colleges: Further results. The Journal of Economic Education 29 (4), 291-300.

Greene, W.H., 2018. Econometric Analysis. Pearson Education, Inc, New York, NY.

Gu, X., Abdel-Aty, M., Xiang, Q., Cai, Q., Yuan, J., 2019. Utilizing UAV video data for in-depth analysis of drivers' crash risk at interchange merging areas. Accident Analysis and Prevention 123, 159-169.

Guevara, C.A., Ben-Akiva, M.E., 2012. Change of scale and forecasting with the control-function method in logit models. Transportation Science 46 (3), 425-437.

Guevara, C.A., 2015. Critical assessment of five methods to correct for endogeneity in discrete-choice models. Transportation Research Part A 82, 240-254.

Guevara, C.A., Hess, S., 2019. A control-function approach to correct for endogeneity in discrete choice models estimated on SP-off-RP data and contrasts with an earlier FIML approach by Train & Wilson. Transportation Research Part B 123, 224-239.

Guo, Y., Li, Z., Wu, Y., Xu, C., 2018. Exploring unobserved heterogeneity in bicyclists' red-light running behaviors at different crossing facilities. Accident

Analysis and Prevention 115, 118-127.

Guo, Y., Li, Z., Liu, P., Wu, Y., 2019. Modeling correlation and heterogeneity in crash rates by collision types using full bayesian random parameters multivariate Tobit model. Accident Analysis and Prevention 128, 164-174.

Harper, C.D., Hendrickson, C.T., Samaras, C., 2016. Cost and benefit estimates of partially-automated vehicle collision avoidance technologies. Accident Analysis and Prevention 95, 104-115.

Hayward, J.C., 1972. Near-miss determination through use of a scale of danger. Highway Research Record 384, 24-34.

Hensher, D.A., Rose, J.M., Greene, W.H., 2015. Applied Choice Analysis. Second edition. Cambridge University Press, Cambridge.

Heydari, S., Miranda-Moreno, L., Hickford, A.J., 2020. On the causal effect of proximity to school on pedestrian safety at signalized intersections: A heterogeneous endogenous econometric model. Analytic Methods in Accident Research 26, 100115.

Heydari, S., Forrest, M., 2024. Estimating the effect of proximity to school on cyclist safety using a simultaneous-equations model with heterogeneity in covariance to address potential endogeneity. Analytic Methods in Accident Research 41, 100318.

Hong Kong Transport Department, 2020. Annual Transport Digest 2020. The Hong Kong SAR Government, Hong Kong.

Hou, Q., Huo, X., Leng, J., 2020. A correlated random parameters tobit model to analyze the safety effects and temporal instability of factors affecting crash rates. Accident Analysis and Prevention 134, 105326.

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.

Hydén, C., 1987. The development of a method for traffic safety evaluation: The Swedish Traffic Conflicts Technique. PhD Thesis. Lund University.

Intini, P., Berloco, N., Colonna, P., Ranieri, V., Ryeng, E., 2018. Exploring the relationships between drivers' familiarity and two-lane rural road accidents. A multi-level study. Accident Analysis and Prevention 111, 280-296.

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., Mannering, F., 2020. A temporal analysis of driver-injury severities in crashes involving aggressive and non-aggressive driving. Analytic Methods in Accident Research 27, 100128.

Islam, M., Alnawmasi, N., 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., Pande, A., 2020. Analysis of single-vehicle roadway departure crashes on rural curved segments accounting for unobserved heterogeneity. Transportation Research Record 2674 (10), 146-157.

Jiménez, F., Naranjo, J.E., García, F., 2013. An improved method to calculate the time-to-collision of two vehicles. International Journal of Intelligent Transportation Systems Research 11 (1), 34-42.

Johnsson, C., Laureshyn, A., Dagostino, C., 2021. Validation of surrogate measures of safety with a focus on bicyclist-motor vehicle interactions. Accident Analysis and Prevention 153, 106037.

Kazemzadeh, K., Afghari, A.P., 2024. Wriggling in the crowd: An inquiry into the interactions between electric bikes and pedestrians in a shared space. Travel Behaviour and Society 36, 100781.

Kim, J.-K., Ulfarsson, G.F., Shankar, V.N., Mannering, F.L., 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.

Krajewski, R., Bock, J., Kloeker, L., Eckstein, L., 2018. The highD dataset: A drone dataset of naturalistic vehicle trajectories on German highways for validation of highly automated driving systems. 21st International Conference on Intelligent Transportation Systems (ITSC), 2118-2125.

Laureshyn, A., Svensson, Å., Hydén, C., 2010. Evaluation of traffic safety, based on micro-level behavioural data: Theoretical framework and first implementation. Accident Analysis and Prevention 42 (6), 1637-1646.

Lee, C., Abdel-Aty, M., 2005. Comprehensive analysis of vehicle–pedestrian crashes at intersections in Florida. Accident Analysis and Prevention 37 (4), 775-786.

Lee, C., Abdel-Aty, M., 2008. Presence of passengers: Does it increase or reduce driver's crash potential? Accident Analysis and Prevention 40 (5), 1703-1712.

Lee, J., Abdel-Aty, M., Cai, Q., Wang, L., 2018. Effects of emergency medical services times on traffic injury severity: A random effects ordered probit approach. Traffic Injury Prevention 19 (6), 577-581.

Li, L., Jiang, R., He, Z., Chen, X., Zhou, X., 2020. Trajectory data-based traffic flow studies: A revisit. Transportation Research Part C 114, 225-240.

Li, Y., Yamamoto, T., Zhang, G., 2018. The effect of fatigue driving on injury severity considering the endogeneity. Journal of Safety Research 64, 11-19.

Li, Y., Gu, R., Lee, J., Yang, M., Chen, Q., Zhang, Y., 2021. The dynamic tradeoff between safety and efficiency in discretionary lane-changing behavior: A random parameters logit approach with heterogeneity in means and variances. Accident Analysis and Prevention 153, 106036.

Li, Z., Liu, P., Wang, W., Xu, C., 2012. Using support vector machine models for crash injury severity analysis. Accident Analysis and Prevention 45, 478-486.

Li, Z., Liu, P., Wang, W., Xu, C., 2014. Development of a control strategy of variable speed limits to reduce rear-end collision risks near freeway recurrent bottlenecks. IEEE Transactions on Intelligent Transportation Systems 15 (2), 866-877.

Lord, D., Mannering, F., 2010. The statistical analysis of crash-frequency data: A review and assessment of methodological alternatives. Transportation Research Part A 44 (5), 291-305.

Madsen, T.K.O., Lahrmann, H., 2017. Comparison of five bicycle facility designs in signalized intersections using traffic conflict studies. Transportation Research Part F 46, 438-450.

Mallia, L., Lazuras, L., Violani, C., Lucidi, F., 2015. Crash risk and aberrant driving behaviors among bus drivers: The role of personality and attitudes towards traffic safety. Accident Analysis and Prevention 79, 145-151.

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

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

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

Mannering, F., Bhat, C.R., 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.

Marshall, R., Summerskill, S., Lenard, J., 2020. The analysis of UK road traffic accident data and its use in the development of a direct vision standard for trucks in London. International Conference on Applied Human Factors and Ergonomics, 427-439.

Martens, M.H., Fox, M.R.J., 2007. Do familiarity and expectations change perception? Drivers' glances and response to changes. Transportation Research Part F 10

(6), 476-492.

Martensen, H., Dupont, E., 2013. Comparing single vehicle and multivehicle fatal road crashes: A joint analysis of road conditions, time variables and driver characteristics. Accident Analysis and Prevention 60, 466-471.

Meng, F., Sze, N.N., Song, C., Chen, T., Zeng, Y., 2021. Temporal instability of truck volume composition on non-truck-involved crash severity using uncorrelated and correlated grouped random parameters binary logit models with space-time variations. Analytic Methods in Accident Research 31, 100168.

Mergia, W.Y., Eustace, D., Chimba, D., Qumsiyeh, M., 2013. Exploring factors contributing to injury severity at freeway merging and diverging locations in Ohio. Accident Analysis and Prevention 55, 202-210.

Mohammadian, S., Haque, M.M., Zheng, Z., Bhaskar, A., 2021. Integrating safety into the fundamental relations of freeway traffic flows: A conflict-based safety assessment framework. Analytic Methods in Accident Research 32, 100187.

Morgan, A., Mannering, F., 2011. The effects of road-surface conditions, age, and gender on driver-injury severities. Accident Analysis and Prevention 43 (5), 1852-1863.

NHTSA, 2020. Traffic safety facts 2018: A compilation of motor vehicle crash data. National Highway Traffic Safety Administration, Washington, D.C.

NHTSA, 2021. Traffic safety facts 2019: A compilation of motor vehicle crash data. National Highway Traffic Safety Administration, Washington, D.C.

NHTSA, 2022. Traffic safety facts 2020: A compilation of motor vehicle crash data. National Highway Traffic Safety Administration, Washington, D.C.

NHTSA, 2023. Traffic safety facts 2021: A compilation of motor vehicle crash data. National Highway Traffic Safety Administration, Washington, D.C.

NHTSA, 2024. MMUCC guideline: Model minimum uniform crash criteria, 6th edition. National Highway Traffic Safety Administration, Washington, D.C.

O'Neill, B., 2009. Preventing Passenger Vehicle Occupant Injuries by Vehicle

Design—A Historical Perspective from IIHS. Traffic Injury Prevention 10 (2), 113-126.

Oviedo-Trespalacios, O., Afghari, A.P., Haque, M.M., 2020. A hierarchical Bayesian multivariate ordered model of distracted drivers' decision to initiate risk-compensating behaviour. Analytic Methods in Accident Research 26, 100121.

Paleti, R., Eluru, N., Bhat, C.R., 2010. Examining the influence of aggressive driving behavior on driver injury severity in traffic crashes. Accident Analysis and Prevention 42 (6), 1839-1854.

Pantangi, S.S., Fountas, G., Anastasopoulos, P.C., Pierowicz, J., Majka, K., Blatt, A., 2020. Do High Visibility Enforcement programs affect aggressive driving behavior? An empirical analysis using Naturalistic Driving Study data. Accident Analysis and Prevention 138, 105361.

Pantangi, S.S., Fountas, G., Sarwar, M.T., Bhargava, A., Mohan, S.B., Savolainen, P., Anastasopoulos, P.C., 2022. The impact of public–private partnerships for roadway projects on traffic safety: An exploratory empirical analysis of crash frequencies. Analytic Methods in Accident Research 33, 100192.

Park, H., Oh, C., Moon, J., Kim, S., 2018. Development of a lane change risk index using vehicle trajectory data. Accident Analysis and Prevention 110, 1-8.

Paul, M., Ghosh, I., 2021. Development of conflict severity index for safety evaluation of severe crash types at unsignalized intersections under mixed traffic. Safety Science 144, 105432.

Persaud, B.N., Retting, R.A., Lyon, C.A., 2004. Crash reduction following installation of centerline rumble strips on rural two-lane roads. Accident Analysis and Prevention 36 (6), 1073-1079.

Peura, C., Kilch, J.A., Clark, D.E., 2015. Evaluating adverse rural crash outcomes using the NHTSA State Data System. Accident Analysis and Prevention 82, 257-262.

Rana, T.A., Sikder, S., Pinjari, A.R., 2010. Copula-based method for addressing endogeneity in models of severity of traffic crash injuries: Application to two-vehicle crashes. Transportation Research Record 2147, 75-87.

Rezapour, M., Moomen, M., Ksaibati, K., 2019. Ordered logistic models of influencing factors on crash injury severity of single and multiple-vehicle downgrade crashes: A case study in Wyoming. Journal of Safety Research 68, 107-118.

Rifaat, S.M., Tay, R., de Barros, A., 2011. Effect of street pattern on the severity of crashes involving vulnerable road users. Accident Analysis and Prevention 43 (1), 276-283.

Saad, M., Abdel-Aty, M., Lee, J., 2019. Analysis of driving behavior at expressway toll plazas. Transportation Research Part F 61, 163-177.

Sacchi, E., Sayed, T., 2016. Conflict-based safety performance functions for predicting traffic collisions by type. Transportation Research Record 2583, 50-55.

Sarwar, M.T., Fountas, G., Anastasopoulos, P.C., 2017. Simultaneous estimation of discrete outcome and continuous dependent variable equations: A bivariate random effects modeling approach with unrestricted instruments. Analytic Methods in Accident Research 16, 23-34.

Savolainen, P., Mannering, F., 2007. Probabilistic models of motorcyclists' injury severities in single- and multi-vehicle crashes. Accident Analysis and Prevention 39 (5), 955-963.

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

Sayed, T., Zaki, M.H., Autey, J., 2013. Automated safety diagnosis of vehicle–bicycle interactions using computer vision analysis. Safety Science 59, 163-172.

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.

Seraneeprakarn, P., Huang, S., Shankar, V., Mannering, F., Venkataraman, N., Milton, J., 2017. Occupant injury severities in hybrid-vehicle involved crashes: A random parameters approach with heterogeneity in means and variances. Analytic Methods in Accident Research 15, 41-55.

Shaon, M.R.R., Qin, X., 2020. Crash data-based investigation into how injury severity is affected by driver errors. Transportation Research Record 2674 (5), 452-464.

Song, P., Sze, N.N., Zheng, O., Abdel-Aty, M., 2022. Addressing unobserved heterogeneity at road user level for the analysis of conflict risk at tunnel toll plaza: A correlated grouped random parameters logit approach with heterogeneity in means. Analytic Methods in Accident Research 36, 100243.

Song, P., Sze, N.N., Chen, S., Labi, S., 2024a. Correcting for endogeneity of crash type in crash injury severity at highway ramp areas. Accident Analysis and Prevention 208, 107785.

Song, P., Sze, N.N., Guo, J., Zhu, D., 2024b. Temporal transferability assessment of injury severity models for single-vehicle and multi-vehicle crashes at freeway ramp areas accounting for unobserved heterogeneity. Under Review.

Stipancic, J., Zangenehpour, S., Miranda-Moreno, L., Saunier, N., Granie, M.A., 2016. Investigating the gender differences on bicycle-vehicle conflicts at urban intersections using an ordered logit methodology. Accident Analysis and Prevention 97, 19-27.

Summerskill, S., Marshall, R., Cook, S., Lenard, J., Richardson, J., 2016. The use of volumetric projections in Digital Human Modelling software for the identification of Large Goods Vehicle blind spots. Applied Ergonomics 53, 267-280.

Sze, N.N., Wong, S.C., 2007. Diagnostic analysis of the logistic model for pedestrian injury severity in traffic crashes. Accident Analysis and Prevention 39 (6), 1267-1278.

Sze, N.N., Wong, S.C., Chan, W.F., 2008. Traffic crashes at toll plazas in Hong Kong. Proceedings of the Institution of Civil Engineers - Transport 161 (2), 71-76.

Tamakloe, R., Hong, J., Park, D., 2020. A copula-based approach for jointly modeling crash severity and number of vehicles involved in express bus crashes on expressways considering temporal stability of data. Accident Analysis and Prevention 146, 105736.

Tarko, A.P., 2021. A unifying view on traffic conflicts and their connection with crashes. Accident Analysis and Prevention 158, 106187.

Tarko, A.P., 2018. Surrogate Measures of Safety. In: Lord, D., Washington, S. eds. Safe Mobility: Challenges, Methodology and Solutions. Emerald Publishing Limited, Bingley, pp. 383-405.

Train, K.E., 2009. Discrete Choice Methods with Simulation. Second edition. Cambridge University Press, New York.

Tsui, K.L., So, F.L., Sze, N.N., Wong, S.C., Leung, T.F., 2009. Misclassification of injury severity among road casualties in police reports. Accident Analysis and Prevention 41 (1), 84-89.

Tulu, G.S., Washington, S., Haque, M.M., King, M.J., 2015. Investigation of pedestrian crashes on two-way two-lane rural roads in Ethiopia. Accident Analysis and Prevention 78, 118-126.

Uzondu, C., Jamson, S., Lai, F., 2018. Exploratory study involving observation of traffic behaviour and conflicts in Nigeria using the Traffic Conflict Technique. Safety Science 110, 273-284.

Wali, B., Khattak, A.J., Karnowski, T., 2020. The relationship between driving volatility in time to collision and crash-injury severity in a naturalistic driving environment. Analytic Methods in Accident Research 28, 100136.

Wang, C., Xu, C., Dai, Y., 2019. A crash prediction method based on bivariate extreme value theory and video-based vehicle trajectory data. Accident Analysis and Prevention 123, 365-373.

Wang, C., Xie, Y., Huang, H., Liu, P., 2021. A review of surrogate safety measures and their applications in connected and automated vehicles safety modeling. Accident Analysis and Prevention 157, 106157.

Wang, Y., Tu, H., Sze, N.N., Li, H., Ruan, X., 2022. A novel traffic conflict risk measure considering the effect of vehicle weight. Journal of Safety Research 80, 1-13.

Wang, Z., Chen, H., Lu, J.J., 2009. Exploring impacts of factors contributing to injury severity at freeway diverge areas. Transportation Research Record 2102, 43-52.

Ward, J.R., Agamennoni, G., Worrall, S., Bender, A., Nebot, E., 2015. Extending Time to Collision for probabilistic reasoning in general traffic scenarios. Transportation Research Part C 51, 66-82.

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

Weng, J., Meng, Q., Yan, X., 2014. Analysis of work zone rear-end crash risk for different vehicle-following patterns. Accident Analysis and Prevention 72, 449-457.

Weng, J., Du, G., Li, D., Yu, Y., 2018. Time-varying mixed logit model for vehicle merging behavior in work zone merging areas. Accident Analysis and Prevention 117, 328-339.

Wong, S.C., Sze, N.N., Hung, W.T., Loo, B.P.Y., Lo, H.K., 2006. The effects of a traffic guidance scheme for auto-toll lanes on traffic safety at toll plazas. Safety Science 44 (9), 753-770.

Wong, S.C., Sze, N.N., Li, Y.C., 2007. Contributory factors to traffic crashes at signalized intersections in Hong Kong. Accident Analysis and Prevention 39 (6), 1107-1113.

Wong, S.C., Sze, N.N., Loo, B.P.Y., Chow, A.S.Y., Lo, H.K., Hung, W.T., 2012. Performance evaluations of the spiral-marking roundabouts in Hong Kong. Journal of Transportation Engineering 138 (11), 1377-1387.

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

Wu, Q., Zhang, G., Chen, C., Tarefder, R., Wang, H., Wei, H., 2016. Heterogeneous impacts of gender-interpreted contributing factors on driver injury severities in single-vehicle rollover crashes. Accident Analysis and Prevention 94, 28-34.

Wu, Y., Abdel-Aty, M., Zheng, O., Cai, Q., Zhang, S., 2020. Automated safety diagnosis based on unmanned aerial vehicle video and deep learning algorithm. Transportation Research Record 2674 (8), 350-359.

Xin, C., Guo, R., Wang, Z., Lu, Q., Lin, P.-S., 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.

Xing, L., He, J., Abdel-Aty, M., Cai, Q., Li, Y., Zheng, O., 2019. Examining traffic conflicts of up stream toll plaza area using vehicles' trajectory data. Accident Analysis and Prevention 125, 174-187.

Xing, L., He, J., Abdel-Aty, M., Wu, Y., Yuan, J., 2020a. Time-varying Analysis of Traffic Conflicts at the Upstream Approach of Toll Plaza. Accident Analysis and Prevention 141, 105539.

Xing, L., He, J., Li, Y., Wu, Y., Yuan, J., Gu, X., 2020b. Comparison of different models for evaluating vehicle collision risks at upstream diverging area of toll plaza. Accident Analysis and Prevention 135, 105343.

Xu, C., Wang, W., Liu, P., Guo, R., Li, Z., 2014. Using the Bayesian updating approach to improve the spatial and temporal transferability of real-time crash risk prediction models. Transportation Research Part C 38, 167-176.

Xu, Z., Zou, X., Oh, T., Vu, H.L., 2021. Studying freeway merging conflicts using virtual reality technology. Journal of Safety Research 76, 16-29.

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.

Yang, D., Xie, K., Ozbay, K., Zhao, Z., Yang, H., 2021. Copula-based joint modeling of crash count and conflict risk measures with accommodation of mixed count-continuous margins. Analytic Methods in Accident Research 31, 100162.

Yasmin, S., Eluru, N., Pinjari, A.R., 2015. Analyzing the continuum of fatal crashes: A generalized ordered approach. Analytic Methods in Accident Research 7, 1-15.

Yasmin, S., Eluru, N., Haque, M.M., 2022. Addressing endogeneity in modeling speed enforcement, crash risk and crash severity simultaneously. Analytic Methods in Accident Research 36, 100242.

Ye, F., Lord, D., 2014. Comparing three commonly used crash severity models on sample size requirements: Multinomial logit, ordered probit and mixed logit models. Analytic Methods in Accident Research 1, 72-85.

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.

Yu, Q., Zhou, Y., Ayele Atumo, E., Qu, L., Zhang, N., Jiang, X., 2023. Addressing endogeneity between hazardous actions and motorcyclist injury severity by integrating generalized propensity score approach and instrumental variable model. Accident Analysis and Prevention 192, 107297.

Yu, R., Abdel-Aty, M., 2013. Multi-level Bayesian analyses for single- and multi-vehicle freeway crashes. Accident Analysis and Prevention 58, 97-105.

Yuan, C., Li, Y., Huang, H., Wang, S., Sun, Z., Li, Y., 2022. Using traffic flow characteristics to predict real-time conflict risk: A novel method for trajectory data analysis. Analytic Methods in Accident Research 35, 100217.

Yun, M., Zhao, J., Zhao, J., Weng, X., Yang, X., 2017. Impact of in-vehicle navigation information on lane-change behavior in urban expressway diverge segments. Accident Analysis and Prevention 106, 53-66.

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.

Zhang, F., Ji, Y., Lv, H., Ma, X., 2021. Analysis of factors influencing delivery e-bikes' red-light running behavior: A correlated mixed binary logit approach. Accident Analysis and Prevention 152, 105977.

Zhang, J., Li, Z., Pu, Z., Xu, C., 2018. Comparing prediction performance for crash injury severity among various machine learning and statistical methods. IEEE Access 6, 60079-60087.

Zheng, L., Sayed, T., Mannering, F., 2021. Modeling traffic conflicts for use in road safety analysis: A review of analytic methods and future directions. Analytic Methods in Accident Research 29, 100142.

Zheng, O., Abdel-Aty, M., Wu, Y., 2019. UCF-SST automated roadway conflicts identify system (A.R.C.I.S). ***https://github.com/ozheng1993/A-R-C-I-S***.

Appendix

Table A1 Results of temporally unconstrained parameter estimation for single-vehicle crashes for 2016

Variable	Coefficient	t-statistic	Marginal effects		
			No injury	Minor injury	Severe injury
Constant [NI]	0.50	2.25			
Constant [SI]	-2.79	-2.29			
Random parameter (normally distributed)					
Vehicle type					
Car [SI]	-4.14	-1.48	-0.0023	-0.0007	0.0030
Standard deviation	2.49	1.66			
Heterogeneity in the mean of random parameter					
Car [SI]: Aggressive driving	3.16	1.89			
Age					
Below 25 [SI]	1.29	1.94	-0.0051	-0.0025	0.0075
Above 59 [SI]	1.63	2.13	-0.0032	-0.0017	0.0049
Alcohol or drugs					
Driving under the influence of alcohol or drugs [NI]	-0.82	-3.31	-0.0154	0.0144	0.0010
Aberrant driving behavior					
Oversteer [NI]	-0.43	-1.79	-0.0086	0.0082	0.0004
Inattentiveness [NI]	-0.60	-2.42	-0.0112	0.0106	0.0005
Safety belt					
Used [SI]	-2.28	-3.28	0.0153	0.0079	-0.0232
Road surface condition					
Dry [SI]	2.11	1.94	-0.0178	-0.0099	0.0277
Road classification					
Interstate highway [NI]	0.33	2.19	0.0378	-0.0360	-0.0018
Collision type					
Overturned [NI]	-1.25	-4.74	-0.0201	0.0186	0.0016
Ramp type					
Off-ramp [NI]	0.45	1.97	0.0461	-0.0439	-0.0022
On-ramp [NI]	0.78	3.26	0.0506	-0.0484	-0.0022
Model statistics					
Number of observations	993				
Degree of freedom	16				
Log-likelihood at zero (LL(0))	-1090.9220				
Log-likelihood at convergence (LL(β))	-618.2351				
McFadden R^2	0.4333				

Table A2 Results of temporally unconstrained parameter estimation for single-vehicle crashes for 2017

Variable	Coefficient	t-statistic	Marginal effects		
			No injury	Minor injury	Severe injury
Constant [NI]	1.42	11.41			
Constant [SI]	-1.94	-2.03			
Random parameter (normally distributed)					
Road surface condition					
Dry [SI]	-1.73	-1.80	-0.0228	-0.0081	0.0309
Standard deviation	3.25	1.98			
Heterogeneity in the mean of random parameter					
Dry [SI]: Merging lane between on-ramp and off-ramp	3.68	1.59			
Gender					
Female [NI]	-0.38	-2.45	-0.0255	0.0243	0.0012
Alcohol or drugs					
Driving under the influence of alcohol or drugs [NI]	-0.64	-2.46	-0.0109	0.0099	0.0011
Aberrant driving behavior					
Oversteer [NI]	-0.54	-2.42	-0.0127	0.0122	0.0005
Aggressive driving [NI]	-1.01	-3.51	-0.0140	0.0129	0.0010
Safety belt					
Used [SI]	-4.34	-3.16	0.0253	0.0092	-0.0345
Area type					
Rural [NI]	0.22	1.35	0.0148	-0.0139	-0.0009
Terrain					
Flat [MI]	-0.44	-1.71	0.0076	-0.0077	0.0001
Collision type					
Overturned [NI]	-1.30	-4.55	-0.0179	0.0172	0.0008
Ramp type					
Off-ramp [SI]	2.41	2.16	-0.0170	-0.0066	0.0236
Model statistics					
Number of observations	1035				
Degree of freedom	14				
Log-likelihood at zero (LL(0))	-1137.0637				
Log-likelihood at convergence (LL(β))	-629.3213				
McFadden R^2	0.4465				

Table A3 Results of temporally unconstrained parameter estimation for single-vehicle crashes for 2018

Variable	Coefficient	t-statistic	Marginal effects		
			No injury	Minor injury	Severe injury
Constant [NI]	1.05	5.95			
Constant [SI]	-1.73	-3.29			
Random parameter (normally distributed)					
Area type					
Rural [NI]	-2.03	-1.99	0.0099	-0.0097	-0.0001
Standard deviation	2.34	2.04			
Heterogeneity in the mean of random parameter					
Rural [NI]: Used safety belt	3.43	2.60			
Heterogeneity in the variance of random parameter					
Rural [NI]: Female driver	0.66	1.69			
Rural [NI]: Speed limit above 60 mph	-0.56	-1.48			
Alcohol or drugs					
Driving under the influence of alcohol or drugs [NI]	-0.75	-2.60	-0.0112	0.0103	0.0008
Aberrant driving behavior					
Aggressive driving [NI]	-0.74	-2.03	-0.0069	0.0065	0.0004
Safety belt					
Used [SI]	-2.05	-3.96	0.0111	0.0080	-0.0191
Road surface condition					
Dry [NI]	-0.53	-3.12	-0.0488	0.0461	0.0027
Lighting condition					
Dark without streetlights [SI]	1.36	2.71	-0.0058	-0.0061	0.0118
Terrain					
Flat [NI]	0.68	1.99	0.0079	-0.0075	-0.0005
Collision type					
Overturned [NI]	-1.87	-4.38	-0.0180	0.0172	0.0008
Crash location					
On traffic lanes [NI]	0.45	2.60	0.0398	-0.0379	-0.0019
Model statistics					
Number of observations	1142				
Degree of freedom	15				
Log-likelihood at zero (LL(0))	-1254.6152				
Log-likelihood at convergence (LL(β))	-680.6298				
McFadden R^2	0.4575				

Table A4 Results of temporally unconstrained parameter estimation for multi-vehicle crashes for 2016

Variable	Coefficient	t-statistic	Marginal effects		
			No injury	Minor injury	Severe injury
Constant [NI]	1.02	9.02			
Constant [SI]	-3.62	-6.58			
Random parameter (normally distributed)					
Speed Limit					
Above 60 mph [NI]	-1.28	-4.40	-0.0050	0.0049	0.0001
Standard deviation	1.22	1.67			
Heterogeneity in the mean of random parameter					
Above 60 mph [NI]: Involvement of alcohol or drugs for drivers	-1.61	-2.73			
Above 60 mph [NI]: Two vehicles	1.75	3.71			
Above 60 mph [NI]: Clear weather	0.40	1.95			
Gender					
Female driver involved [MI]	0.16	1.74	-0.0172	0.0173	-0.0001
Maneuver					
Making maneuver action [SI]	-1.15	-1.89	0.0017	0.0008	-0.0025
Truck					
Truck involved [SI]	1.52	2.49	-0.0009	-0.0005	0.0014
Lighting condition					
Dusk or dawn [SI]	1.65	2.11	-0.0005	-0.0002	0.0008
Dark with streetlights [NI]	-0.45	-3.29	-0.0070	0.0069	0.0001
Area type					
Rural [NI]	-0.29	-2.92	-0.0110	0.0109	0.0002
Road classification					
Interstate highway [NI]	0.23	2.27	0.0180	-0.0177	-0.0003
Road configuration					
One-way mainline [NI]	0.47	4.98	0.0268	-0.0265	-0.0003
Collision type					
Sideswipe collision [NI]	0.84	6.51	0.0193	-0.0190	-0.0004
Angle collision [NI]	-0.40	-3.33	-0.0106	0.0104	0.0002
Ramp type					
On-ramp [MI]	0.16	1.72	-0.0065	0.0065	0.0000
Model statistics					
Number of observations	4172				
Degree of freedom	18				
Log-likelihood at zero (LL(0))	-4583.4105				
Log-likelihood at convergence (LL(β))	-2154.5966				
McFadden R^2	0.5299				

Parameter defined for: [NI] No injury; [MI] Minor Injury; [SI] Severe Injury

Table A5 Results of temporally unconstrained parameter estimation for multi-vehicle crashes for 2017

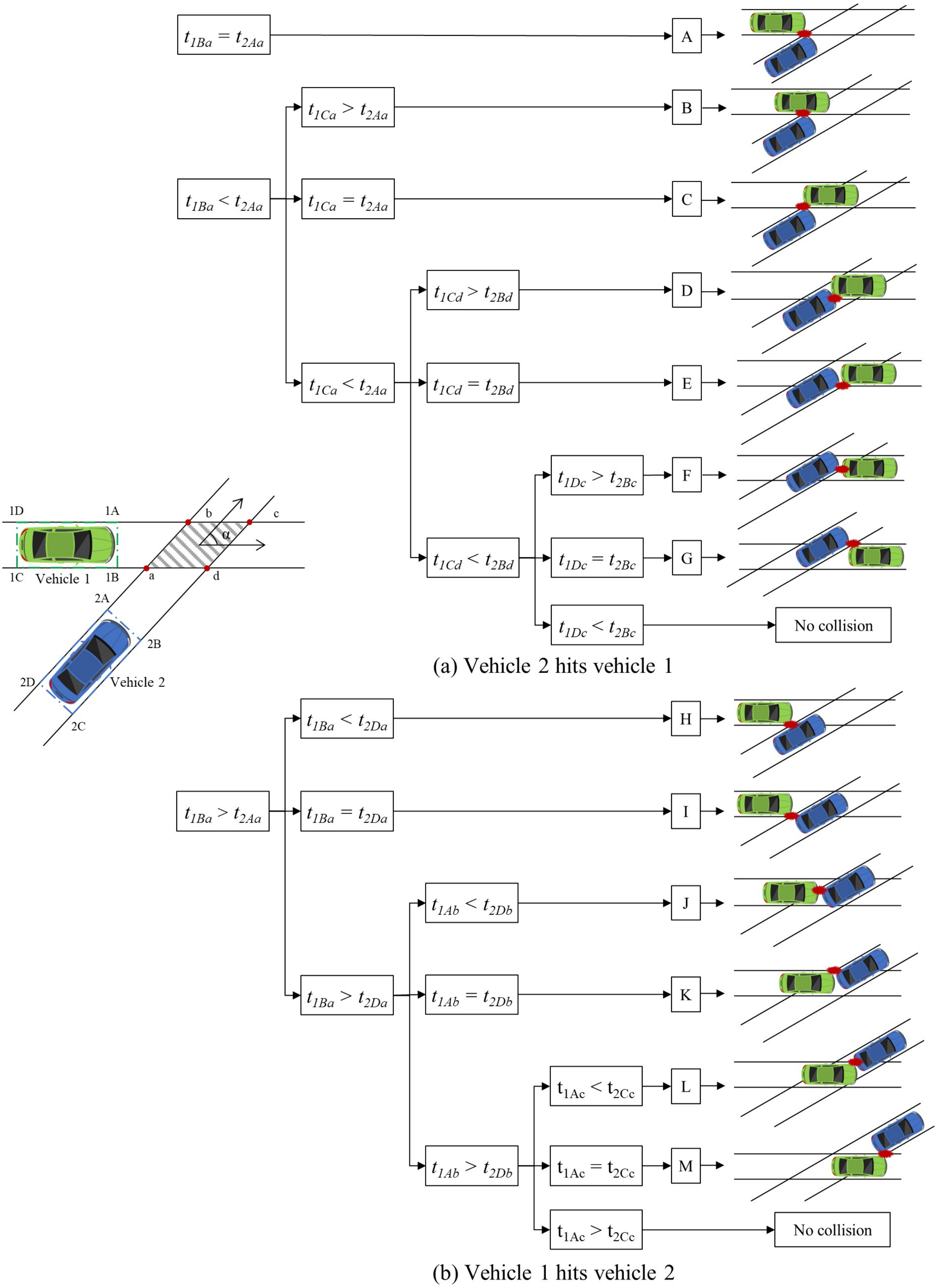
Variable	Coefficient	t-statistic	Marginal effects		
			No injury	Minor injury	Severe injury
Constant [NI]	1.38	11.49			
Constant [SI]	-1.63	-2.82			
Random parameter (normally distributed)					
Lighting condition					
Dark with streetlights [NI]	-2.26	-1.56	-0.0041	0.0040	0.0001
Standard deviation	4.27	1.89			
Speed limit					
Above 60 mph [NI]	-1.81	-2.25	-0.0165	0.0157	0.0007
Standard deviation	3.09	3.57			
Heterogeneity in the mean of random parameter					
Dark with streetlights [NI]: Two vehicles	4.26	1.68			
Above 60 mph [NI]: Truck involved	-0.59	-1.81			
Above 60 mph [NI]: Two vehicles	1.92	3.44			
Above 60 mph [NI]: On traffic lanes	1.34	1.76			
Heterogeneity in the variance of random parameter					
Above 60 mph [NI]: Making maneuver action	-0.31	-1.87			
Gender					
Female driver involved [NI]	-0.51	-4.69	-0.0476	0.0464	0.0012
Alcohol or drugs					
Involvement of alcohol or drugs for drivers [NI]	-1.71	-3.93	-0.0034	0.0033	0.0001
Maneuver					
Making maneuver action [SI]	-1.11	-2.49	0.0026	0.0020	-0.0046
Truck					
Truck involved [SI]	1.53	3.59	-0.0017	-0.0014	0.0031
Number of vehicles involved					
Two vehicles [SI]	-1.55	-3.28	0.0038	0.0028	-0.0067
Lighting condition					
Dusk or dawn [NI]	-0.58	-2.81	-0.0035	0.0034	0.0001
Road configuration					
One-way mainline [NI]	0.56	5.29	0.0260	-0.0254	-0.0006
Collision type					
Sideswipe collision [NI]	1.19	7.60	0.0192	-0.0185	-0.0007
Angle collision [NI]	-0.51	-3.93	-0.0108	0.0106	0.0003
Ramp type					
On-ramp [MI]	0.21	1.93	-0.0062	0.0064	-0.0002
Model statistics					
Number of observations	4491				
Degree of freedom	21				
Log-likelihood at zero (LL(0))	-4933.8678				
Log-likelihood at convergence (LL(β))	-2322.1116				
McFadden R^2	0.5294				

Parameter defined for: [NI] No injury; [MI] Minor Injury; [SI] Severe Injury

Table A6 Results of temporally unconstrained parameter estimation for multi-vehicle crashes for 2018

Variable	Coefficient	t-statistic	Marginal effects		
			No injury	Minor injury	Severe injury
Constant [NI]	1.49	14.1			
Constant [SI]	-1.78	-3.01			
Random parameter (normally distributed)					
Speed Limit					
Above 60 mph [MI]	0.77	2.85	-0.0066	0.0068	-0.0002
Standard deviation	1.41	1.85			
Heterogeneity in the mean of random parameter					
Above 60 mph [MI]: Truck involved	0.69	2.79			
Above 60 mph [MI]: Two vehicles	-1.74	-5.01			
Heterogeneity in the variance of random parameter					
Above 60 mph [MI]: Straight horizontal alignment	-0.70	-2.49			
Above 60 mph [MI]: Rear-end collision	0.78	1.76			
Gender					
Female driver involved [NI]	-0.37	-4.01	-0.0358	0.0348	0.0010
Alcohol or drugs					
Involvement of alcohol or drugs for drivers [NI]	-1.19	-4.21	-0.0034	0.0032	0.0001
Maneuver					
Making maneuver action [SI]	-1.27	-2.97	0.0034	0.0010	-0.0043
Truck					
Truck involved [SI]	2.09	5.12	-0.0035	-0.0010	0.0045
Number of vehicles involved					
Two vehicles [SI]	-1.54	-2.99	0.0052	0.0013	-0.0065
Lighting condition					
Dark without streetlights [NI]	-0.29	-1.98	-0.0036	0.0035	0.0001
Area type					
Rural [NI]	-0.18	-1.92	-0.0073	0.0070	0.0003
Road configuration					
One-way mainline [NI]	0.29	3.17	0.0159	-0.0155	-0.0004
Collision type					
Sideswipe collision [NI]	1.04	7.7	0.0190	-0.0181	-0.0009
Angle collision [NI]	-0.68	-5.7	-0.0153	0.0149	0.0004
Model statistics					
Number of observations	4878				
Degree of freedom	18				
Log-likelihood at zero (LL(0))	-5359.0307				
Log-likelihood at convergence (LL(β))	-2376.8405				
McFadden R^2	0.5565				

Parameter defined for: [NI] No injury; [MI] Minor Injury; [SI] Severe Injury



(a) Vehicle 2 hits vehicle 1

(b) Vehicle 1 hits vehicle 2

Figure A1 Illustration of possible conflict scenarios

if $t_{1BA} < t_{2Aa}$ **then**
  if $t_{1Ca} < t_{2Aa}$ **then**
    if $t_{1Cd} < t_{2Bd}$ **then**
      if $t_{1Dc} > t_{2Bc}$ **then**
        $\mathrm{TTC} = \frac{t_{2Bd} \times t_{1Dc} - t_{1Cd} \times t_{2Bc}}{t_{2Bd} + t_{1Dc} - t_{1Cd} - t_{2Bc}}$
      else if $t_{1Dc} = t_{2Bc}$ **then**
        $\mathrm{TTC} = t_{2Bc}$ (or t_{1Dc})
    else if $t_{1Cd} > t_{2Bd}$ **then**
      $\mathrm{TTC} = \frac{t_{1Cd} \times t_{2Aa} - t_{2Bd} \times t_{1Ca}}{t_{1Cd} + t_{2Aa} - t_{2Bd} - t_{1Ca}}$
    else
      $\mathrm{TTC} = t_{2Bd}$ (or t_{1Cd})
  else if $t_{1Ca} > t_{2Aa}$ **then**
    $\mathrm{TTC} = t_{2Aa}$
  else
    $\mathrm{TTC} = t_{2Aa}$ (or t_{1Ca})
else if $t_{1BA} > t_{2Aa}$ **then**
  if $t_{1Ba} > t_{2Da}$ **then**
    if $t_{1Ab} > t_{2Db}$ **then**
      if $t_{1Ac} < t_{2Cc}$ **then**
        $\mathrm{TTC} = \frac{t_{2Cc} \times t_{1Ab} - t_{1Ac} \times t_{2Db}}{t_{2Cc} + t_{1Ab} - t_{1Ac} - t_{2Db}}$
      else if $t_{1Ac} = t_{2Cc}$ **then**
        $\mathrm{TTC} = t_{1Ac}$ (or t_{2Cc})
    else if $t_{1Ab} < t_{2Db}$ **then**
      $\mathrm{TTC} = \frac{t_{2Db} \times t_{1Ba} - t_{1Ab} \times t_{2Da}}{t_{2Db} + t_{1Ba} - t_{1Ab} - t_{2Da}}$
    else
      $\mathrm{TTC} = t_{1Ab}$ (or t_{2Db})
  else if $t_{1Ba} < t_{2Da}$ **then**
    $\mathrm{TTC} = t_{1Ba}$
  else
    $\mathrm{TTC} = t_{1Ba}$ (or t_{2Da})
else
    $\mathrm{TTC} = t_{1BA}$ (or t_{2Aa})

Figure A2 Formulation for modified TTC