University of Memphis
University of Memphis Digital Commons

Electronic Theses and Dissertations

12-3-2016

A Comprehensive Discrete Choice Analysis of Injury Severity in Roadway Work Zone Crashes

Mohamed Osman

Follow this and additional works at: https://digitalcommons.memphis.edu/etd

Recommended Citation
Osman, Mohamed, "A Comprehensive Discrete Choice Analysis of Injury Severity in Roadway Work Zone Crashes" (2016). *Electronic Theses and Dissertations*. 1568.
https://digitalcommons.memphis.edu/etd/1568

This Dissertation is brought to you for free and open access by University of Memphis Digital Commons. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of University of Memphis Digital Commons. For more information, please contact khggerty@memphis.edu.

A COMPREHENSIVE DISCRETE CHOICE ANALYSIS OF INJURY SEVERITY IN ROADWAY WORK ZONE CRASHES

by

Mohamed G. Osman

A Dissertation

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

Major: Civil Engineering

The University of Memphis

December 2016

Copyright© Mohamed G. Osman

All rights reserved

DEDICATION

To my parents, my lovely wife, and my beautiful kids for their love and support.

ACKNOWLEDGEMENTS

I would like to express my innermost appreciation to a number of individuals who, after the blessing of God Almighty, helped making this work possible.

First, I would like to thank my mentor and dissertation committee chair Dr. Sabya Mishra for all of his invaluable knowledge, time, expertise, management skills, support, and so much more throughout my graduate-level studies. I feel privileged to have met Dr. Mishra (University of Memphis) and to have had the opportunity to be under his direction on the path of accomplishing this work and ultimately obtaining my Ph.D. Degree. Not only have I learned academically from Dr. Mishra, he had introduced me to the world of publications; an achievement that was beyond my thoughts at the beginning of my graduate course work. I would like to also thank Dr. Mishra for the financial support through providing GAUSS 15; a powerful statistical package utilized in the analysis conducted in this dissertation.

I am grateful to have worked with my dissertation advisory committee members: Dr. Mihalis Gkolias (University of Memphis), Dr. Stephanie Ivey (University of Memphis), Dr. Rajesh Paleti (Old Dominion University), and Dr. Tao Ma (University of Memphis).

I am especially thankful to Dr. Stephanie Ivey for her support and guidance during the course-work choices made at the beginning stages of my Ph.D. Degree. I am thankful for her time during several discussions regarding a variety of dissertation topics. I have gained tremendous knowledge from Dr. Ivey during several graduate and undergraduate courses. My decision to pursue a Ph.D. Degree was, in large part, a result of her constant encouragement throughout my time at the University of Memphis. Truly, a wonderful person!

I would like to thank Dr. Mihalis M. Gkolias for the wealth of knowledge I have gained from him during several graduate courses, a number of past and on-going research studies and publications and during the preparation of this dissertation document. Dr. Gkolias has always been there for me throughout my graduate studies which I am grateful for.

I sincerely acknowledge the support of Dr. Rajesh Paleti during the development and scripting of the econometric frameworks mostly utilized in this dissertation. Dr. Paleti has provided invaluable comments throughout the chapters forming this dissertation, which helped enrich its content. I am thankful to Dr. Paleti for always providing time to discuss, explain, and resolve issues encountered during the progress of this dissertation. Dr. Paleti's contributions helped adding value to publications served as parts in this dissertation.

I would like to Thank Dr. Tao Ma for his time and assistance in reviewing the work accomplished in this dissertation. I am thankful to Dr. Ma for being a valuable member of this dissertation advisory committee. Dr. Ma has been graciously proactive offering support and assistance during various stages of completing this dissertation. Dr. Ma has provided significant comments that increased the value and quality of the dissertation.

I would like to express my sincere appreciation to the faculty and staff in the Civil Engineering Department at the University of Memphis for each and every piece of information I have learned and the knowledge I have gained over the years.

I am grateful to the State of Tennessee and the Tennessee Department of Transportation (TDOT) as my employer to have allowed me as a Professional Engineer to utilize the State's Fee Waiver Program to cover approximately fifty percent of the tuition costs of my course work during my graduate studies from 2007 through 2016.

Lastly, I am forever thankful to my family; without their constant love and support, I wouldn't be where I am today. I sincerely praise God Almighty for giving me the strength and companionship in every step of the way.

ABSTRACT

Osman, Mohamed G. PhD. The University of Memphis. December, 2016. A Comprehensive Discrete Choice Analysis of Injury Severity in Roadway Work Zone Crashes. Major Professor: Dr. Sabyasachee Mishra.

Work zones are critical parts of the transportation infrastructure renewal process consisting of rehabilitation of roadways, maintenance, and utility work. Given the specific nature of a work zone (complex arrangements of traffic control devices and signs, narrow lanes, duration) a number of crashes occur with varying severities involving different vehicle sizes.

This dissertation proposes a comprehensive discrete choice analysis of injury severity of crashes in work zones on both the crash and occupant levels, in roadway work zones through a comprehensive set of discrete choice econometric frameworks. Robust discrete choice modeling structures are introduced and applied in the field of work zone safety.

This dissertation contains three (3) studies representing the empirical analysis conducted to address the following research questions:

1. What factors may contribute to the injury severity levels of large-truck crashes in work zones? And what are the robust analytical methods to recognize such factors?
2. How do specific work zone configurations affect factors contributing to the levels of injury severity of work zone crashes?
3. How does the specific work zone-component-area where a crash has occurred affect factors contributing to the injury severity levels of work zone crashes?

The first study investigates the causal factors contributing to injury severity of large truck crashes in work zones. The second study investigates the causal factors contributing to the injury severity of passenger-car occupants for crashes occurring in different work zone configurations (lane closure, lane shift/crossover, shoulder/median, intermittent, and other). The third study investigates the causal factors contributing to driver's injury severity in the different work zone component-areas (advance-warning, transition, activity, and termination areas). The first study

compares a comprehensive set of discrete choice modeling structures; Multinomial Logit (MNL) model, Nested Logit (NL) model, Ordered Logit (ORL) model and Generalized Ordered Response Logit (GORL) model. The second and third studies developed the Mixed Generalized Ordered Response Probit (MGORP) modeling framework to conduct the proposed analysis to answer the second and third research questions. The empirical analysis was conducted using work zone crash database in 10 years of the Highway Safety Information System (HSIS).

PREFACE

This dissertation is submitted for the degree of Doctor of Philosophy at the University of Memphis. The research described herein was conducted under the supervision of Dr. Sabyasachee Mishra, Department of Civil Engineering, University of Memphis, between January 2014 and December 2016.

This work is, to the best of my knowledge, original, except where acknowledgements and references are made to previous work. To my knowledge, neither this dissertation nor any other substantially similar dissertation has been or is being submitted for any other degree at any other university.

Part of this work has been presented in the following publication:

Osman, M., Paleti, R., Mishra, S., Golias, M.M., 2016. Analysis of injury severity of large truck crashes in work zones. Accid. Anal. Prev. 97, 261–273. doi:10.1016/j.aap.2016.10.020

Osman, M., Paleti, R., Mishra, S., Golias, M.M., 2016. Analysis of injury severity of large truck crashes in work zones. Presented at the Transportation Research Board 95th Annual Meeting, Washington, D.C.

Another part of this work has been accepted for presentation at the 96th Annual Board Meeting of the Transportation Research Board (TRB).

Osman, M., Paleti, R., Mishra, S., Golias, M.M., 2017. Analysis of Driver Injury Severity for Different Work Zone Areas. To be presented at the Transportation Research Board 96th Annual Meeting, Washington, D.C.

TABLE OF CONTENTS

Chapter Page

DEDICATION III

ACKNOWLEDGEMENTS IV

ABSTRACT VI

PREFACE VIII

TABLE OF CONTENTS IX

LIST OF TABLES XIII

LIST OF FIGURES XIV

1. INTRODUCTION 1

1.1. Contributions 2

1.2. Structure of the Manuscript 4

2. LITERATURE REVIEW 6

2.1. Injury Severity of Large Truck Crashes 6

2.1.1. Large Truck Crash Severity 7

2.1.2. Work Zone Crash Severity 8

2.1.3. Large Truck Crash Severity in Work Zones 9

2.2. Injury Severity of Crashes for Different Work Zone Configurations 10

2.3. Injury Severity of Crashes for Specific Work Zone Component-Areas 10

2.4. Literature Review Summary 11

3. METHODOLOGY 13

3.1. Econometric Framework 13

3.1.1. Unordered Modeling Framework 14
3.1.1.1. Multinomial logit (MNL) model 15
3.1.1.2. Nested logit (NL) model 15
3.1.2. Ordered Response Framework 16
3.1.2.1 Ordered logit (ORL) model 16
3.1.2.2. Generalized ordered logit (GORL) model 17
3.1.2.3. Mixed generalized ordered response probit (MGORP) model 18
4. MODEL ESTIMATION AND ANALYSIS OF INJURY SEVERITY 21
4.1. Injury Severity of Large Truck Crashes in Work Zones 21
4.1.1. Introduction 21
4.1.2. Data 22
4.1.3. Empirical Analysis 23
4.1.4. Estimation Results 25
4.1.4.1. Roadway characteristics 26
4.1.4.2. Traffic characteristics 27
4.1.4.3. Environmental characteristics 31
4.1.4.4. Temporal characteristics 31
4.1.4.5. Work zone characteristics 32
4.1.4.6. Crash characteristics 33
4.1.5. In-Depth Analysis of Large Truck Exposure over Time 35
4.1.6. Measures of Fit 37
4.1.7. Elasticity Effects 37
4.1.7.1. Elasticity effects of GORL model 38
4.1.8. Conclusions 40
4.2. Injury Severity of passenger-Car Crashes for Different Work Zone Configurations 43

4.2.1. Introduction 43
4.2.2. Data 45
4.2.3. Empirical Analysis 46
4.2.4. Estimation Results 49
4.2.4.1. Roadway characteristics 50
4.2.4.2. Environmental characteristics 54
4.2.4.3. Traffic characteristics 55
4.2.4.4. Work Zone characteristics 55
4.2.4.5. Temporal characteristics 57
4.2.4.6. Crash characteristics 58
4.2.3. Measures of Fit 59
4.2.4. Elasticity Effects 59
4.2.4.1. Elasticity effects of MGORP model 60
4.2.5. Implications of Variable Effects and Recommendations 64
4.2.5.1. Recommendations for all work zone configurations 64
4.2.5.2. Recommendations for specific work zone configurations 65
4.2.6. Conclusions 66
4.3. Injury Severity of drivers for Different Work Zone component-Areas 69
4.3.1. Introduction 69
4.3.2. Data 70
4.3.3. Empirical Analysis 71
4.3.4. Estimation Results 74
4.3.4.1. Roadway characteristics 75
4.3.4.2. Environmental characteristics 75
4.3.4.3. Traffic characteristics 76
4.3.4.4. Work zone characteristics 76

4.3.4.5. Temporal characteristics 79
4.3.4.6. Crash characteristics 79
4.3.4.7. Driver attributes 80
4.3.4.8. Vehicle characteristics 81
4.3.5. Measures of Fit 83
4.3.6. Elasticity Effects 83
4.3.6.1. Elasticity effects of MGORP model 84
4.3.7. Conclusions 88
5. CONCLUSIONS AND FUTURE RESEARCH 90
5.1. Conclusions 90
5.2. Avenues for Future Research 91
REFERENCES 93

LIST OF TABLES

TABLE 1 Initial frequency of dependent variable (large truck) 22

TABLE 2 Final frequency of dependent variable (large truck) 22

TABLE 3 Frequency distribution of explanatory variables (large truck) 24

TABLE 4 MNL, ORL, GORL model results 28

TABLE 5 GORL-time-of-day interactions model results 35

TABLE 6 Elasticity effects of GORL 39

TABLE 7 Initial frequency of dependent variable (passenger-car) 46

TABLE 8 Final frequency of dependent variable (passenger-car) 46

TABLE 9 Frequency distribution of explanatory variable (passenger-car) 47

TABLE 10 MGORP model results (passenger-car) 51

TABLE 11 Elasticity effects of MGORP model for "Severe Injury" outcome (passenger-car) 62

TABLE 12 Initial frequency of dependent variable (driver) 71

TABLE 13 Final frequency of dependent variable (driver) 71

TABLE 14 Frequency distribution of explanatory variable (driver) 72

TABLE 15 MGORP model results (driver) 77

TABLE 16 Elasticity effects of MGORP model for "Severe Injury" outcome (driver) 85

LIST OF FIGURES

FIGURE 1 Growth in VMT roadway lane miles 1

FIGURE 2 Work zone configurations 44

FIGURE 3 Crash frequency distribution by work zone configuration 49

FIGURE 4 Work zone component-areas 69

FIGURE 5 Crash frequency distribution by work zone component-area 73

1. INTRODUCTION

Work zone safety is a major concern for the Federal Highway Administration (FHWA), State Departments of Transportation (DOTs), and the public. Fig. 1 indicates that over the last 30 years, the total lane miles in the US have nearly increased by 7.4% whereas the Vehicle Miles Travelled (VMT) increased by 86% (FHWA, 2012). With increased VMT, work zone fatalities and injuries have also increased. Nationally, there were 87,606 work zone crashes in 2010 which is approximately 1.6% of the total number of roadway crashes. More than 20,000 workers were injured in work zones in 2010. In the same year, work zone crashes resulted in 37,476 injuries which equates to approximately four injuries every hour. In 2010, there were 514 fatal crashes resulting in 576 fatalities in work zones, which equates to approximately one fatality every 15 hours (FHWA, 2010). Work zones have unique traffic conditions that are different from other crash locations and thus warrant studies that focus exclusively on these locations instead of generally pooling them with other locations.

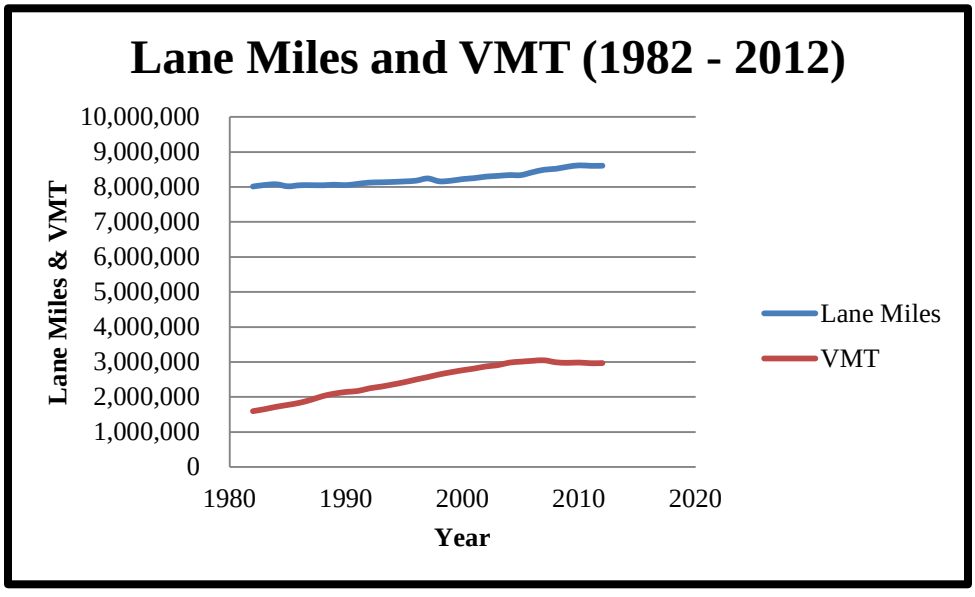


Fig. 1. Growth in VMT roadway lane miles

Source of data developing graph:
http://www.fhwa.dot.gov/policyinformation/statistics/2012/vmt422c.cfm

Additionally, according to (FHWA), in more recent years, the number of work zone-related crashes has been declining, following a nationally similar decreasing trend in highway crashes (FHWA, 2016). However, in 2013 alone, the number of work zone-related crashes was nationally estimated to be 67,523 (FHWA, 2016). In 2013, despite the downward annual trend in the number of work zone crashes, the number of work zone injuries has increased (FHWA, 2016). Approximately 47,758 non-fatal injuries were reported to have occurred in work zones in 2013 (FHWA, 2016). In the same year, there were 527 fatal crashes in work zones resulting in 579 fatalities (FHWA, 2016).

Another key segment of crashes, that is of major concern both to the transportation officials and the trucking industry, are those involving large trucks. In 2012 alone, there were 317,000 large truck crashes in the US that resulted in 3,464 fatalities and 73,000 injuries (FHWA, 2014). In the same year, large trucks accounted for 8% of all vehicles involved in fatal crashes and 3% of vehicles involved in injury and property-damage-only (PDO) crashes (U.S. Department of Transportation, 2014). Although these percentages may not seem alarming at first glance, the economic impact could be substantial because large truck crashes incur high costs including high value goods, and higher travel delays associated with longer traffic incident durations. Moreover, the determinants of the injury severity level of crashes involving large trucks can be considerably different from crashes involving passenger cars and/or relatively smaller commercial fleet. So, it is important to focus exclusively on large truck crashes in order to be able to understand the relative effect of the different factors on the injury severity levels of those individuals involved. Although, on average, 85% of fatalities in work zones were drivers or occupants of passenger cars (FHWA, 2016).

1.1. Contributions

Part of this dissertation aims to contribute to the literature on work zone safety by exploring the characteristics of large truck crashes in work zones using a disaggregate-level

analytical approach that focusses on each individual crash and associated set of potentially contributing factors. Specifically, the study examines the factors that impact the severity level of the most severely injured individual involved in the crash, which essentially marks the overall severity level of the crash. Understanding large truck crash severity characteristics in work zones will be a steppingstone in enabling practitioners, designers, and DOT officials to mitigate the severity of such type of crash. The findings of this study have important implications in the work zone safety field, education of motorists, training of truck drivers, and traffic regulation and control. Designers of roadway work zones will be able to implement effective safety measures that will allow DOT officials to better manage the safety of a work zone through learning about the important factors influencing crashes involving large trucks.

According to the work zone safety literature, there have not been any studies that undertook analysis at the level of the specific work zone configuration where a crash has occurred. Depending on the nature of the temporary traffic control (TTC) plan pertaining to a specific work zone configuration, the determinants and the magnitude of impact of factors that influence injury severity of crashes that occur in work zones can vary across different work zone configurations. Another part of this dissertation aims to address this gap in the literature by developing an analytical model of the injury severity of the most injured passenger-car occupant in work zone crashes by exploring interactions between the different work zone configurations and the potential associated risk factors. Understanding the different characteristics contributing to the injury severity of passenger-car crashes in the different work zone configurations will serve as a great advantage enabling practitioners, designers, and DOT officials to mitigate the severity of those individuals; generally involved in a work zone crash or particularly within a specific work zone configuration. As stated in the 2009 edition of the Manual on Uniform Traffic Control Devices (MUTCD), TTC applications were designed as minimum solutions for the depicted configurations (“Manual on Uniform Traffic Control Devices (MUTCD),” 2009). Therefore, work zone designers and DOTs can make informed decision when upgrading TTC plans from

those minimums to best suit their needs by possessing advanced knowledge of what factors may or may not affect the injury severity levels of motorists based on the work zone configuration it is.

According to the 2009 edition of the MUTCD, a work zone comprises of four areas: (1) advance warning area, (2) transition area, (3) activity area, and (4) termination area. Each of these areas has a specific purpose and may vary in size and location depending on the nature of the work activity. Very few studies undertook analysis at the level of the specific work zone component-area where a crash has occurred. Therefore, another part of this dissertation aims to address this gap in the literature by developing an analytical model of driver injury severity in work zone crashes by exploring the interactions between the above identified work zone component-areas and the potential associated risk factors. Understanding the different characteristics affecting the severity of driver in the different work zone areas will serve as a great advantage enabling practitioners, designers, and DOT officials to mitigate the severity of those individuals generally involved in a work zone crash or particularly within specific component-areas of the work zone.

1.2. Structure of the Manuscript

The structure of the rest of this dissertation is as follows: Chapter 2 presents a comprehensive literature review of the related studies. The first subsection of the literature review mainly focuses on the injury severity of crashes specifically involving large trucks, occurring generally in work zones, and those that focus on both large truck crash severity and work zone safety combined. The following subsection of the literature review specifically discusses past studies related to the injury severity of crashes occurring in different work zone configurations. The last subsection of the literature review focuses on the research that is related to the specific work zone component-area where a crash has occurred in.

Chapter 3 presents the methodology followed in this research to include the different econometric frameworks and statistical modeling techniques developed to analyze the injury severity of crashes in work zones. To be specific, this chapter describes each of the discrete choice econometric modeling utilized in this research. Additionally, chapter 3 introduces state-of-the-art modeling structures developed specifically to address limitation of previous models generally utilized in the injury severity past literature.

Chapter 4 of this dissertation presents model estimations for different work zones settings with different vehicles sizes to include both crash-level as well as occupant-level analysis. Specifically, three (3) studies are conducted to include several model estimations to address each of the three research questions proposed in this research. For each of the three studies conducted, chapter 4 also presents the different datasets utilized, a detailed description and frequency distributions of both dependent and independent variables entered the modeling process. Chapter 4 also describes the empirical analysis, modeling estimations, and a detailed interpretation of the estimation results for all proposed models within each of the three (3) proposed studies. Finally, chapter 4 presents measures-of-fit within each study and elasticity effects of variables entered each proposed model.

Chapter 5 of this dissertation presents final conclusions, recommendations, limitations of each study, and finally provides avenues for future work.

2. LITERATURE REVIEW

The overview of the injury severity literature indicates that there is a vast body of research generally pertaining to studying crash injury severity. Due to the broad nature of such a safety topic, and based on the specific studies conducted in this dissertation, the literature review section chapter subcategorizes the injury severity literature into the following subsections: (1) injury severity of large truck crashes, (2) injury severity of passenger-car crashes for different work zone configurations, and (3) injury severity of drivers in for different work zone component-areas.

2.1. Injury Severity of Large Truck Crashes

Several research studies have been conducted to analyze the severity of crashes involving large trucks (Chang and Mannering, 1999; Duncan et al., 1998; Islam and Hernandez, 2013; Li and Bai, 2009; Pahukula et al., 2015; Qi et al., 2013; Wang and Shi, 2013; Wang et al., 2010). The overview of the literature indicates that there is a vast body of research examining the factors affecting the severity of large truck-involved accidents on both crash-level and occupant-level. The literature presented in this subsection is primarily focused on injury severity of large trucks in work zones at the crash-level to obtain insights and to help to meet the goal of this research. However, occupant-level injury severity studies are imperative in the context of work zone safety and comprehensively presented in the literature (Chang and Chien, 2013; Chen and Chen, 2011; Dong et al., 2015; Khorashadi et al., 2005; Lemp et al., 2011; Mooradian et al., 2013; Wong et al., 2011; Zhu and Srinivasan, 2011a, 2011b).

The past literature can be grouped under three categories – (1) those that focus exclusively on large truck crash severity modeling, (2) those that focus on injury severity in the context of work zone safety, and (3) those that focus both on large truck crash severity and work zone safety combined. In this section we present a review of the crash-level literature that specifically pertained to injury severity of crashes involving large trucks, work zones, or both.

The econometric framework comparisons utilized in this study have been recently used by other researchers in the context of injury severity analysis to evaluate alternate discrete outcome frameworks for modeling crash injury severity (Yasmin and Eluru, 2013). Sample size requirements were evaluated by comparing three commonly crash severity models (Ye and Lord, 2014). Another study has evaluated alternate discrete choice frameworks for modeling ordinal discrete variables but not necessarily in the context of injury severity (Eluru, 2013). A discrete choice model comparison was applied to investigate cyclist injury severity in automobile-involved bicycle crashes (Chen and Shen, 2016). Pedestrian Injury Severity in New York City was also examined using alternative ordered response frameworks (Yasmin et al., 2014). To our knowledge, this is the first application of such a comprehensive set of discrete choice models in the context of work zone safety. A brief overview of past literature in these three categories follows in the next three subsections.

2.1.1. Large Truck Crash Severity

A variety of discrete choice models were used in the literature to analyze large truck crash severity. For example, assessing the severity of truck crashes on a freeway network using a hierarchical regression model indicated that the presence of ramp, freeway segment length, and weather conditions were important factors affecting truck safety performance (Wang and Shi, 2013). Utilizing nested logit models to investigate the severity in truck and non-truck crashes, risk factors that are unique to large trucks were identified. Variables that increased injury severity for large trucks were higher speed limits, vehicles making right or left turns, and rear-end collisions (Chang and Mannering, 1999). Using a random-parameter ordered probit model allowed the identification of the differences between random and fixed factors affecting the severity outcome. It was found that the severity level is highly influenced by complex interactions between factors, and that the effects of some variables can vary across observations (Islam and Hernandez, 2013). Investigating rear-end large truck crashes using an ordered probit model indicated that darkness,

high speed differential between vehicles and trucks, higher speed limits, wet surfaces on a grade, a car struck to the rear, and alcohol increased crash severity while snow and ice, congested roads, and station wagon decreased the likelihood of a severe crash (Duncan et al., 1998). An exploratory study utilized a mixed logit model to analyze injury severity of crashes involving large trucks on Texas highways which revealed that time-of-day (12-6 AM), summer time (June-August), clear weather, rural areas, and 4-lane roadways were all contributing factors to higher likelihood of higher injury severity levels (M. Islam and Hernandez, 2013). Another study also used mixed logit models to estimate the effect of time of day on injury severity of large truck crashes in urban areas (Pahukula et al., 2015). The study uncovered major differences both in the combination of variables and their magnitude of impact on the severity outcomes across different time periods. Among different explanatory variables used in the study, the effects of traffic flow, lighting road surface conditions, time of year, and percentage of trucks were found to vary by time period (Pahukula et al., 2015). In recent years, mixed logit models have generally gained attention within the discrete choice modeling literature due to their flexibility in allowing variations over data observations as compared the restrictions imposed by standard logit models. This modeling technique has been utilized in previous large truck literature, but not necessarily within the context of injury severity (Romo et al., 2014).

2.1.2. Work Zone Crash Severity

A work zone crash is defined as a crash that occurred in an area comprising a work zone as per defined by the Manual of Uniform Traffic Control Devices (MUTCD). Specifically, for the purpose of this study, a work zone extends from the “advanced warning area” until the “termination area”. There is some literature that focused specifically on crashes in work zones. For instance, one study used the ordered probit model to analyze severity of rear-end crashes in work zones. The study found that alcohol, night hours, pedestrians, roadway defects, truck-involvement, and the number of vehicles involved increased crash severity, while careless

backing, stalled vehicles, slippery surfaces, and misunderstanding flagging signals resulted in less severe injuries in the event of a crash (Qi et al., 2013). However, there is no consensus on these findings in the safety literature. Other studies that used similar discrete choice modeling methods found slightly contradicting results (Wang et al., 2010). Another study by Wong et al., 2011) examined factors influencing injury severity of highway workers in work zone intrusion crashes using multiple correspondence analysis, Cox proportional hazard regression, logistic regression, and Poisson regression models and found that work zone location and duration, time of the day, and type of activity performed by workers were the most significant factors impacting severity outcomes.

2.1.3. Large Truck Crash Severity in Work Zones

Most of the crash severity literature to date provide only basic information in terms of the large truck-involvement in a work zone crash (Li and Bai, 2008a; Qi et al., 2013; Wang et al., 2010). Such studies used large truck-involvement in a work zone crash as a binary explanatory variable in severity models. There is only one study in the literature that modeled injury severity of crashes involving large trucks in work zones. Khattak and Targa, 2004 have modeled injury severity and total harm in work zone crashes involving large trucks by assigning an economic cost for the different severity levels (Khattak and Targa, 2004). The study found that, on average, large truck crashes that occurred on two-way undivided roads, roads with higher speed limits, and in the proximity of work zones tend to be more severe than other crashes. Given the relatively sparse literature on work zone crashes involving large trucks, the current study aims to develop improved tools that can provide better insights by using new econometric methods that were developed recently. Specifically, the current study compared the performance of alternate modeling frameworks in identifying significant factors affecting the severity of large truck crashes in work zones.

2.2. Injury Severity of Crashes for Different Work Zone Configurations

Within the work zone crash severity literature, some studies mainly focused on fatal crashes (Arditi et al., 2007; Daniel et al., 2000; Schrock et al., 2004), other studies discussed on both fatal and injury crashes (Elghamrawy et al., n.d.; Li and Bai, 2008b), and some conducted injury severity analyses (Akepati and Dissanayake, 2011; Khattak et al., 2002; Khattak and Targa, 2004; Li and Bai, 2009; Qi et al., 2013; Wang et al., 2010). There have been inconsistencies in the literature regarding whether work zone crashes are more severe relative to those occurring in non-work zone areas. Some studies indicated that work zone crashes were in fact more severe (Bédard et al., 2002; Garber and Zhao, 2002; Meng et al., 2010; Pigman and Agent, 1990; Ullman et al., 2006), while others disagreed (FHWA, 2016; Hargroves and Martin, 1980; Nemeth and Migletz, 1978; Nemeth and Rathi, 1983; Rouphail et al., 1988). According to the work zone safety literature, there have not been any studies that undertook analysis at the level of the specific work zone configuration where a crash has occurred.

2.3. Injury Severity of Crashes for Specific Work Zone Component-Areas

Earlier studies on work zone safety focused on different aspects including crash risk factors, severity, type, location, rate, and time frame. Due to the broad nature of these past studies, this subsection of the literature review will mainly focus on studies related to work zone crash severity and risk factors. Very few studies undertook analysis at the level of the specific work zone area where a crash has occurred. One previous study analyzed the distribution and characteristics of crashes in specific areas within a work zone and compared selected characteristics of work zone crashes with those of non-work zone crashes (Garber and Zhao, 2002). This study concluded that the activity area was more susceptible to crashes regardless of road type while the termination area had the lowest frequency. Also, the study found that most nighttime work zone crashes were in the activity area and that the severity of crashes in the

daytime and night time were not significantly different. However, the study did not develop an analytical model of injury severity in different work zone areas.

2.4. Literature Review Summary

As a result of the conducted literature review, the following gaps in the most recent and robust research venues can be outlined:

a. In terms of the different econometric modeling frameworks utilized in the literature to analyze injury severity outcomes, a wide variety of discrete choice modeling were adopted; although there have not been a conclusive evidence of the superiority of one model compared to another. A comparison of discreet choice modeling frameworks represents a critical need; generally for injury severity research, and particularly within the work zone safety field.

b. In terms of injury severity analyses, the majority of the up-to-date researches conducted have analyzed the contributing factors to a specific crash injury severity outcome generally in work zones, involving large trucks while very few studies have analyzed both categories combined. Within the work zone safety field, it is crucial to understand and clearly specify the most contributing factors leading to specific injury severity levels of crashes involving large trucks specifically occurring within the work zone area. The findings of this study have important implications in the work zone safety field, education of motorists, training of truck drivers, and traffic regulation and control. Designers of roadway work zones will be able to implement effective safety measures that will allow DOT officials to better manage the safety of a work zone through learning about the important factors influencing crashes involving large trucks.

c. Different work zone configurations can influence causal factors contributing to the degree of injury severity of passenger-car crashes. Due to the specific work zone layouts and different applications of TTC plans, some work zone configurations may increase the

risk of injury severity of a passenger-car crash, while other configurations may reduce this risk. The literature is sparse in terms of analyzing different work zone configurations. This study aims to fill in this gap in the literature.

d. Another gap in the was found; that is specifically analyzing the casual factors leading to certain injury literature severity outcomes as a result of a crash occurring in a particular work zone component-area. As previously mentioned, only one previous study analyzed the distribution and characteristics of crashes in specific areas within a work zone and compared selected characteristics of work zone crashes with those of non-work zone crashes (Garber and Zhao, 2002). Understanding the different characteristics affecting the severity of driver in the different work zone areas will enable practitioners, designers, and DOT officials to mitigate the severity of those individuals generally involved in a work zone crash or particularly within specific component-areas of the work zone.

3. METHODOLOGY

The methodology chapter of this dissertation describes the different econometric frameworks utilized in the three (3) studies conducted to answer the three proposed research questions. The first study in this research proposes a model comparison between the most common discrete choice modeling frameworks in the injury severity literature in order to be able to distinguish the best-fit modeling structure among the rest of the proposed models to analyze injury severity data in work zones. Aside from models that have been extensively utilized in the literature for the past three decades, this study also proposes more innovative econometric frameworks that address most of the limitations of the previous models.

3.1. Econometric Framework

The modeling methods typically used to analyze crash data pertaining injury severity can be grouped into two categories – unordered (Chang and Mannering, 1999; Holdridge et al., 2005; Savolainen and Mannering, 2007; Shankar et al., 1996; Ulfarsson and Mannering, 2004) and ordered (Eluru et al., 2008; Wang et al., 2010; Zhu and Srinivasan, 2011b). In the unordered modeling framework, the observed severity outcome is assumed to be the outcome with the highest latent severity function value (there is one severity function corresponding to each severity outcome). Each of the latent severity functions is specified as a linear function of different crash factors with a stochastic component to account for all unobserved factors that influence the corresponding severity outcome. The coefficients in all the severity functions constitute the set of parameters that are estimated using inference methods such as the maximum likelihood (ML) estimation approach. In the ordered framework, on the other hand, a single latent propensity function is assumed to be translated into the observed severity outcome depending on the value of the propensity function relative to threshold parameters (number of thresholds = number of possible severity outcomes – 1). The latent propensity function is specified as a function of different factors along with a stochastic component to account for all

unobserved factors that influence crash severity. The parameters in the single propensity equation and the thresholds constitute the set of parameters that are estimated using methods such as the maximum likelihood (ML). Earlier comparison studies for analyzing ordinal discrete outcomes (not necessarily in the context of severity analysis) found that the unordered framework fits data better than ordinal models because of the flexibility provided by additional parameters in the unordered models. However, a study by Eluru *et al*., (2008) developed generalized ordered models that allow parameterization of the threshold parameters providing additional flexibility to the ordinal models (Eluru et al., 2008). So, it is not surprising that a recent comparison analysis of unordered and ordered frameworks that considers generalized version of ordered models found minor differences between the two models (Anowar et al., 2014). So, it is imperative that researchers compare and choose the best method specific to the empirical context of interest. This section describes the two modeling frameworks and their generalized variants used in this study.

3.1.1. Unordered Modeling Framework

Let *i* be the index for the injury severity outcome (1 = "no injury", 2 = "injury", and 3 = "severe injury") and *n* be the index for crash. Also, let *I* denote the total number of severity outcomes (which is 3 in the current empirical context) and *N* denote the total number of crashes in the dataset. In this study, a linear-in-parameter specification was adopted for the deterministic part of U_{in} as follows: $U_{in} = \boldsymbol{\beta}_i' \boldsymbol{X}_{in} + \varepsilon_{in}$ where $\boldsymbol{X}_{in}$ is a $K_i \times 1$ vector of exogenous covariates (including crash factors, work zone attributes, environmental, and roadway conditions), $\boldsymbol{\beta}_i$ is the corresponding $K_i \times 1$ vector of coefficients and ε_{in} denotes all the unobserved factors that influence the severity function for outcome *i* in crash *n*. As discussed earlier, in the unordered framework, the observed severity outcome is the severity outcome with the highest latent severity function value. So, the probability that crash *n* sustains severity outcome *i*, $P_n(i)$ is given by Equation (1):

$$P_n(i) = P\left(\boldsymbol{\beta}_i' \boldsymbol{X}_{in} + \varepsilon_{in} \geq \boldsymbol{\beta}_j' \boldsymbol{X}_{jn} + \varepsilon_{jn}\right) \forall\, j \neq i \quad (1)$$

3.1.1.1. Multinomial logit (MNL) model

In the MNL model, the stochastic components ε_{in} in the latent severity functions U_{in} are assumed to be independent and identically distributed (*i.i.d.*) across different severity outcomes and crashes. Moreover, the identical distribution is assumed to be standard type-1 extreme value distribution (also referred to as Gumbel distribution). Given these assumptions on the stochastic term ε_{in}, $P_n(i)$ can be derived to be Equation (2):

$$P_n(i) = \frac{\exp(\boldsymbol{\beta}_i' \boldsymbol{X}_{in})}{\sum_{\forall I} \exp(\boldsymbol{\beta}_j' \boldsymbol{X}_{jn})} \quad (2)$$

The $\sum_{i=1}^{I} K_i$ parameters in the MNL model were estimated by maximizing the log-likelihood function obtained by taking the natural logarithm of the product of probabilities of observed severity outcomes given by Equation (3) as follows:

$$LL = \sum_{n=1}^{N} \left(\sum_{i=1}^{I} \delta_{in} \right) \quad (3)$$

where δ_{in} is defined as 1 if the observed severity outcome for crash n is i and zero otherwise.

3.1.1.2. Nested logit (NL) model

The MNL model has the Independence of Irrelevant Alternatives (IIA) property which implies that changes in conditions that influence one severity outcome do not change the relative probabilities of other severity outcomes. This can be a strong restrictive assumption in the current empirical context given that severity data is ordinal in nature with potentially strong correlations between successive severity outcomes. Past literature found evidence for correlation among unobserved effects to be present (Shankar et al., 1996), while other research has not (Shankar and Mannering, 1996). Assuming the IIA property to hold in cases when it is violated can produce incorrect parameter estimates because of specification errors. The NL model that relaxes the IIA assumption by allowing correlation in unobserved factors of subsets of alternatives is more suited for such scenarios (Shankar et al., 1996). In this study, alternate two-level nesting structures that

group all the severity outcomes into *S* mutually exclusive and exhaustive nests B_s each with nesting parameter ρ_s $(0 < \rho_s \leq 1)$ were estimated. The probability of severity outcome *i* that belongs to nest B_r can be obtained as the product of conditional probability of the outcome *i* within the nest B_r and the probability of the nest B_r among all possible nests B_s $s \in \{1,2, \ldots . S\}$ as in Equation (4):

$$P_n(i) = \frac{e^{\frac{\boldsymbol{\beta}_i' \boldsymbol{X}_{in}}{\rho_r}}}{\sum_{k \in B_r} e^{\frac{\boldsymbol{\beta}_k' \boldsymbol{X}_{kn}}{\rho_r}}} \times \frac{e^{\rho_r IV_r}}{\sum_{s=1}^{S} e^{\rho_s IV_s}} \text{ where Inclusive Value } (IV_s) = LN\left[\sum_{k \in B_s} e^{\frac{\boldsymbol{\beta}_k' \boldsymbol{X}_{kn}}{\rho_s}}\right] \quad (4)$$

3.1.2. Ordered Response Framework

3.1.2.1 Ordered logit (ORL) model

As discussed earlier, in the ordinal framework, latent propensity y_n^* is translated into observed severity outcomes by threshold parameters. This study adopted a linear-in-parameter specification for the observed part of y_n^* and a standard logistic distribution that is *i.i.d.* across crashes for the stochastic component ε_n . The equation system for the ORL model is shown as Equation (5) (McKelvey and Zavoina, 1975):

$$\begin{aligned} y_n^* &= \boldsymbol{\beta}' \boldsymbol{X}_n + \varepsilon_n \\ P_n(i) &= P(\psi_{i-1} < y_n^* < \psi_i) \\ &= P(\psi_{i-1} < \boldsymbol{\beta}' \boldsymbol{X}_n + \varepsilon_n < \psi_i) \\ &= P(\psi_{i-1} - \boldsymbol{\beta}' \boldsymbol{X}_n < \varepsilon_n < \psi_i - \boldsymbol{\beta}' \boldsymbol{X}_n) \\ &= F(\psi_i - \boldsymbol{\beta}' \boldsymbol{X}_n) - F(\psi_{i-1} - \boldsymbol{\beta}' \boldsymbol{X}_n) \end{aligned} \quad (5)$$

where $\boldsymbol{X}_n$ is $K \times 1$ vector of covariates and $\boldsymbol{\beta}$ is the corresponding $K \times 1$ vector of coefficients; $\psi_i' s$ are threshold parameters; $\psi_0 = -\infty \; and \; \psi_{I+1} = \infty$; $F(.)$ is the standard logistic cumulative distribution function. The model structure requires that the thresholds to be strictly ordered for the partitioning of the latent risk propensity measure into the ordered injury severity

categories($i.e., -\infty < \psi_1 < \psi_2 < \cdots < \psi_{I-1} < \infty$). The parameters in the ORL model ($\boldsymbol{\beta}$ and $\psi_i's$) were estimated using the ML inference method.

3.1.2.2. Generalized ordered logit (GORL) model

One of the restrictive assumptions of the standard ORL model is that it assumes that the threshold parameters do not vary across different crashes. Eluru *et al.*(2008) relaxed this assumption by parameterizing the thresholds as a function of exogenous factors providing additional flexibility to the model (Eluru et al., 2008). The structure of the GORL in Equation (6) follows the same structure of the ORL in Equation (5) except for ψ parameters which are now subscripted by index n to reflect that these parameters will vary across crashes (Eluru et al., 2008; Romo et al., 2014).

$$y_n^* = \boldsymbol{\beta}'\boldsymbol{X}_n + \varepsilon_n$$

$$P_n(i) = P\left(\psi_{n,i-1} < y_n^* < \psi_{n,i}\right)$$

$$= F\left(\psi_{n,i} - \boldsymbol{\beta}'\boldsymbol{X}_n\right) - F\left(\psi_{n,i-1} - \boldsymbol{\beta}'\boldsymbol{X}_n\right) \quad (6)$$

To ensure strict ordering of thresholds, the parameterization in Equation (7) was adopted:

$$\psi_{n,i} = \psi_{n,i-1} + exp(\alpha_i + \boldsymbol{\gamma}_i'\boldsymbol{Z}_{ni}) \quad (7)$$

where z_{ni} is a set of exogenous variables associated with the i^{th} threshold excluding the constant; $\boldsymbol{\gamma}_i$ is the corresponding vector of coefficients, and α_i is a parameter associated with injury severity level $i = 1,2,\ldots,I-1$. $\psi_{n,1}$ is specified as $exp(\alpha_1)$ for identification reasons. The ORL model can be obtained from the GORL model by imposing the constraints that $\gamma_i = 0$ for all i.

Another aspect of considerable importance in injury severity analysis is unobserved heterogeneity. Injury severity conditional on crash occurrence can depend on numerous factors all of which are most certainly not observed in crash databases. These unobserved factors can moderate the influence of other observed covariates in the model leading to variation in the

parameter effects across different observations. This unobserved variation in covariate effects is referred to as *unobserved heterogeneity.* Mannering *et al.* (2016) describes this issue in greater detail and presented alternate modeling methods available in the literature for handling the problem (Mannering et al., 2016). Among these methods, the random parameters methods are the most prominent. Consistent with the recommendations of this study, the random parameters or Mixed GOR Probit (MGORP) model for analyzing injury severity of most-injured passenger car occupant involved in work zone crashes was adopted. A brief overview of the MGORP model follows.

3.1.2.3. Mixed generalized ordered response probit (MGORP) model

Let $n(n = 1,2, \dots, N)$ be an index that represents crashes and $i(i = 1,2, \dots, I)$ is the index representing injury severity categories. In the context of this study, index i will take the value "no injury" $(i = 1)$, "injury" $(i = 2)$, and "severe injury" $(i = 3)$. The MGORP model starts as a standard ORP. The equation system for the ORP model is shown by Equation (8) (McKelvey and Zavoina, 1975):

$$y_n^* = \boldsymbol{\beta}' \boldsymbol{X}_n + \varepsilon_n$$
$$y_n = i \; if \; (\psi_{i-1} < y_n^* < \psi_i) \tag{8}$$

where y_n^* is the latent propensity for most injured occupant in crash n, which is translated into observed severity outcomes y_n by threshold parameters ψ_i. $\mathbf{X}_n$ is $K \times 1$ vector of covariates and $\boldsymbol{\beta}$ is the corresponding $K \times 1$ vector of coefficients; $\psi_i's$ are threshold parameters; $\psi_0 = -\infty$ and $\psi_{I+1} = \infty$. ε_n is a random error term capturing the effects of unobserved factors on the injury severity propensity. For model identification purposes, this error term ε_n is assumed to be independently and identically standard normal distributed across the crashes which leads to the ordered probit model (ORP). The model structure requires that the thresholds to be strictly

ordered for the partitioning of the latent risk propensity measure into the ordered injury severity categories (i. e. , $-\infty < \psi_1 < \psi_2 < \cdots < \psi_{I-1} < \infty$) for each crash n.

The enhancement of the ORP model to a MGORP is characterized by the enabling $\boldsymbol{\beta}$ vector and ψ thresholds to vary across observations. This is accomplished through subscripting these parameters with the index n. The MGORP equation system can then be written as in Equation (9):

$$y_n^* = \boldsymbol{\beta}'_n \boldsymbol{X}_n + \varepsilon_n$$
$$y_n = i \; if \; \left(\psi_{n,i-1} < y_n^* < \psi_{n,i}\right) \tag{9}$$

To account for unobserved heterogeneity, the $\boldsymbol{\beta}_n$ vector is assumed to a realization from a multivariate normal distribution with mean $\boldsymbol{\beta}$ and covariance $\boldsymbol{\Sigma}$. Now, Equation (9) can be re-written as in Equation (10):

$$y_n^* = \boldsymbol{\beta}_n \boldsymbol{X}_n + \tilde{\varepsilon}_n \text{ where } \tilde{\varepsilon}_n \sim N(0, \boldsymbol{X}'_n \boldsymbol{\Sigma} \boldsymbol{X}_n)$$
$$y_n = i \; if \; \left(\psi_{n,i-1} < y_n^* < \psi_{n,i}\right) \tag{10}$$

Also, Equation (11) shows that a specific non-linear functional form was used for parameterizing thresholds to ensure that the ordinal criterion is met ($-\infty < \psi_{n,1} < \psi_{n,2} < \cdots < \psi_{n,I-1} < \infty$) for each crash n:

$$\psi_{n,i} = \psi_{n,i-1} + exp\left(\alpha_{n,i} + \boldsymbol{\gamma}'_{n,i} \boldsymbol{Z}_{ni}\right) \tag{11}$$

where $\boldsymbol{Z}_{ni}$ is a set of exogenous variables associated with the i^{th} threshold excluding the constant; $\boldsymbol{\gamma}_{n,i}$is the corresponding vector of coefficients, and $\alpha_{n,i}$ is a parameter associated with injury severity level $i = 1,2, \ldots, I-1$. $\psi_{n,1}$ is specified as $exp(\alpha_1)$ for identification reasons. Moreover, $\boldsymbol{\gamma}_{n,i}$ vector is assumed a realization from a multivariate normal distribution with mean

$\boldsymbol{\gamma}_i$ and covariance $\boldsymbol{\Omega}_i$. Let $\boldsymbol{\gamma}_n$ and $\boldsymbol{\gamma}$ be the vertically stacked column vectors of all $\boldsymbol{\gamma}_{ni}$ and $\boldsymbol{\gamma}_i$ vectors.

The probability of observed injury severity *i* of crash *n* conditional on $\boldsymbol{\gamma}_n$ is given by Equation (12):

$$P_n(i|\boldsymbol{\gamma}_n) = \Phi\left(\frac{\psi_{n,i} - \boldsymbol{\beta}'_n \boldsymbol{X}_n}{\sqrt{\boldsymbol{X}'_n \boldsymbol{\Sigma} \boldsymbol{X}_n}}\right) - \Phi\left(\frac{\psi_{n,i-1} - \boldsymbol{\beta}'_n \boldsymbol{X}_n}{\sqrt{\boldsymbol{X}'_n \boldsymbol{\Sigma} \boldsymbol{X}_n}}\right) \quad (12)$$

The unconditional probability can be obtained by integrating out the random components of $\boldsymbol{\gamma}_n$ using simulation. Within the studies utilized the MGORP in this research, the resulting models' parameters were estimated using the maximum simulated likelihood (MSL) inference approach and 150 Halton draws (Bhat, 2001).

4. MODEL ESTIMATION AND ANALYSIS OF INJURY SEVERITY

4.1. Injury Severity of Large Truck Crashes in Work Zones

4.1.1. Introduction

The current study aims to contribute to the literature on work zone safety by exploring the characteristics of large truck crashes in work zones using a disaggregate-level analytical approach that focusses on each individual crash and associated set of potentially contributing factors. Specifically, the study examines the factors that impact the severity level of the most severely injured individual involved in the crash, which essentially marks the overall severity level of the crash. Understanding large truck crash severity characteristics in work zones will be a steppingstone in enabling practitioners, designers, and DOT officials to mitigate the severity of such type of crash. The findings of this study have important implications in the work zone safety field, education of motorists, training of truck drivers, and traffic regulation and control. Designers of roadway work zones will be able to implement effective safety measures that will allow DOT officials to better manage the safety of a work zone through learning about the important factors influencing crashes involving large trucks.

Considering the discrete nature of injury severity categories, a number of comparable econometric models were developed including multinomial logit (MNL), nested logit (NL), ordered logit (ORL), and generalized ordered logit (GORL) models. The MNL and NL models belong to the class of unordered discrete choice models and do not recognize the intrinsic ordinal nature of the injury severity data. The ORL and GORL models, on the other hand, belong to the ordered response framework that was specifically developed for handling ordinal dependent variables. Past literature did not find conclusive evidence in support of either framework. This study aims to compare these alternate modeling frameworks in order to find the best-fit modeling structure for generally analyzing injury severity crash data. To fulfil the goals of this study, the model comparison is conducted for analyzing injury severity of crashes involving large trucks in

work zones. The model estimation utilizes a compiled database of crashes that (1) involved large trucks and (2) occurred in work zones in the past 10 years in Minnesota. Empirical findings indicate that the GORL model provided superior data fit as compared to all

4.1.2. Data

A dataset consisting of work zone crashes over 10 years (2003-2012) in Minnesota (MN) was collected from the Highway Safety Information System (HSIS). Two main datasets were obtained and merged. The first was the "accident file", containing variables such as crash time, location, roadway condition, crash type, traffic control, and weather conditions. The second was the "road file", containing basic characteristics of the roadway segment where the crash occurred such as lane, shoulder and median widths, speed limit, and several geometric design variables. For the purposes of this study, only crashes involving at least one large truck were considered as truck-related crashes. The dataset contained 18,889 crashes in work zones with 15% involving large trucks (*i.e.*, 2,881 records were available for the analysis in this study). The crash severity level followed the KABCO injury severity scale where K=killed, A=incapacitating injury,

TABLE 1 **Initial frequency of dependent variable (large truck)**

Injury Severity Category	**Count**	**(%)**
Fatal (K)	19	0.66%
Incapacitating Injury (A)	29	1.01%
Non-Incapacitating Injury (B)	152	5.28%
Possible Injury (C)	435	15.10%
Property Damage (O)	2,246	77.96%
Total	2,881	100.00%

TABLE 2 **Final frequency of dependent variable (large truck)**

Combined Injury Severity Category	**Count**	**(%)**
Severe Injury (K,A,B)	200	6.94%
Injury (C)	435	15.10%
No Injury (O)	2,246	77.96%
Total	2,881	100.00%

B=non-incapacitating injury, C=possible injury, and O=no injury. The distribution of crashes by injury severity is presented in Tables 1 and 2. Table 1 shows the percentage of each severity category of the original data. Due to the low frequency of some of the severity levels, some of the severity categories were combined. The combined injury severity categories are shown in Table 2. Fatal, incapacitating, and non-incapacitating severity levels were combined into one severity level called "severe injury". "Possible injury" which is referred to as "injury" and "no injury" categories were kept as is.

4.1.3. Empirical Analysis

Several categories of independent variables were considered in the empirical analysis to account for roadway, traffic, environmental, temporal, work zone, and crash characteristics. Table 3 indicates the frequency distribution of the explanatory variables. Roadway characteristics included functional class and geometric design factors. Functional class of each roadway was classified into one of the following types - "rural principal arterial", "urban principal arterial", "urban minor arterial", and "collectors, local systems or rural minor arterial". Geometric design factors included whether the road was curved or straight, number of lanes, and whether the roadway was curbed and access-controlled. Traffic characteristics included "speed limit" upstream of a work zone area. The effect of speed was captured using three categorical variables indicating whether speed limit was less than 35 mph, between 35 and 40 mph, between 45 and 50 mph, between 55 and 60 mph, or greater than 60 mph. Work zone immediate upstream speed limits were utilized in this research for each crash location. Environmental factors included wet surface and adverse weather (rain, fog, and snow). The impact of time of day was captured using three broad time categories - day (6 am - 6 pm), evening (6 pm - 12 am), and late night (12 am - 6 am). In addition to the time-of-day variables, an indicator variable for peak hours that denoted whether the crash occurred between 7-10 am or 4-7 pm was used. Work zone characteristics included the type of work zone (lane closure, shoulder or median work, lane shift or crossover, and intermittent/moving work zones). The crash work zone location indicated whether the crash

TABLE 3 Frequency distribution of explanatory variables (large truck)

Explanatory Variable	(%)	Explanatory Variable	(%)
Roadway		Time of the day	
Functional class		Day (6:00 AM - 6:00 PM)	6.4
Rural principal arterial	13.0	Evening (6:00 PM - 12:00 AM)	81.1
Urban principal arterial	57.7	Late night (12:00 AM - 6:00 AM)	12.5
Urban minor arterial	15.3	**Work zone**	
Other	14.0	Workers present	
Geometric design		Yes	42.4
Alignment		No	57.6
Curved	15.3	Work zone type	
Straight	84.7	Lane closure	36.6
Number of lanes		Lane shift/crossover	19.1
Two-lane	19.9	Shoulder or median	20.3
Multi-lane	80.1	Intermittent/moving	7.6
Curb		Other	16.4
Yes	37.5	Work zone location	
No	62.5	Advanced signs	8.0
Access control		Transition	18.8
Full	52.3	Activity	53.0
Partial	6.6	Termination	2.6
None	41.1	Other	17.6
Traffic		**Crash**	
Speed limit (mph)		Number of vehicles	
< 35	21.2	Single-vehicle	12.0
35 - 40	6.8	Multi-vehicle	88.0
45 – 50	12.4	Truck type	
55 - 60	50.7	Bus	7.9
65 - 70	8.9	2 axle 1 unit	15.6
Environmental		3+ axle 1 unit	11.7
Roadway surface condition		1 unit with trailer	6.5
Wet	15.8	Tractor-semitrailer	48.5
Dry	84.2	Other	9.8
Weather condition		Location	
Adverse	8.7	Signalized intersection	
Clear	91.3	Yes	15.1
Temporal		No	84.9
Peak hours		On-bridge	
Peak	34.7	Yes	6.5
Off-peak	65.3	No	93.5

has occurred in the proximity of advanced signage, work activity, transition, or termination areas. In addition to the variables listed above, an indicator variable for whether workers were present at the work zone was also tested during model estimation. Several geometric design variables were

purposely omitted, such as (lane width, median width, shoulder widths), due to the fact that those types of variables are, most of the time, altered in a work zone depending on the nature and type of roadway work it is. This level of detailed work zone-specific geometric layout data was not available to the authors. Lastly, crash characteristics included the number of vehicles involved in the crash, truck type, and whether the crash occurred at a signalized intersection or on a bridge. The final specifications for the presented models were based on a logical process of removing the statistically insignificant variables and combining other variables when their effects were statistically insignificant. The model estimation process was, in large part, guided by findings of past research and intuitiveness of the parameters estimated. Since work zones are naturally different than regular roadway segments in terms of roadway geometry, traffic controls and operational characteristics, the injury severity results in the current analysis are considered to be distinctive for work zones due to the special characteristics of roadways in work zones versus non-work zone areas. The final sample in the current study was narrowed down to those accidents that only occurred in a work zone while involving at least one large truck.

4.1.4. Estimation Results

Table 4 presents the estimation results of the MNL, ORL, and GORL models. To test the validity of the IIA assumption of the MNL model, two-level nested logit (NL) models with two possible nesting structures with three severity outcomes were estimated. Neither nesting structure was found to be statistically sound as both nesting parameters did not fall between 0 and 1 (Manski and McFadden, 1981). So, the NL model was excluded from further analysis. The results corresponding to the MNL model consists of two columns labelled “injury”, and “severe injury”, while “no injury” category was chosen as the base category. The ORL model has one column corresponding to the variables in the propensity specification and two threshold parameters. The results corresponding to the GORL model are presented in two columns; the first column corresponds to the variables in the latent risk propensity (not including a constant) and the second

column corresponds to the variables in the second threshold specification between the "injury" and "severe injury" outcomes. The respective t-values of the estimated coefficients are shown in parentheses. Table 4 also presents the initial log-likelihood value, the log-likelihood value at convergence, the Bayesian information criterion value (BIC), the McFaddenR^2, and the total number of crashes n for the three models.

4.1.4.1. Roadway characteristics

Rural principal arterials increased the likelihood of "severe injury" relative to "no injury" outcomes according to the MNL. Similar results were obtained from the ORL and GORL models. However, other functional class categories were also found to be statistically significant in the ordered response framework. To be specific, the OR models indicate that, on average, rural principal, urban principal, and minor urban arterials have higher risk propensity relative to rural minor arterials, collectors, and local systems.

Curved roadways were found to be associated with lower likelihood of sustaining "injury" but higher likelihood of "severe injury" relative to "no injury" outcomes in the event of a crash. This non-monotonic effect of road curvature is interesting. In some cases, it seems that steep curves are dangerous and can lead to severe outcomes in the event of crash and in few other cases, increased driver awareness and cautious driving while maneuvering curved roads reduces chances of injury (Lemp et al., 2011).

Crashes on two-lane roadways tend to be less severe compared to crashes on multi-lanes roads. This finding is contrary to other studies that found that work zone crashes on two-lane roads were more severe; however these earlier studies focused on all crashes in work zones and did not control for the presence of a large truck (Li and Bai, 2009; Wang et al., 2010). All three models showed that crashes in work zones of curbed roadways were less severe compared to crashes on non-curbed roadways. It is important to note that, unlike in the ORL model in which the variable was present in the propensity equation, this variable was found to influence injury

severity through the threshold parameter between the “injury” and “severe injury” outcomes. Specifically, a positive coefficient for ‘curbed’ roadway in the GORL threshold specification suggests wider translation region or higher likelihood of “injury” versus lower likelihood of “severe injury” outcomes, in the event of a crash.

Lack of access-control increased the likelihood of “injury” and “severe injury” relative “no injury” outcomes according to the MNL model. The positive coefficient values for the ORL and the GORL latent propensities showed similar results. Non-access-controlled roadways are likely to have more conflict points. The negative coefficient value for non-access-controlled roadways in the threshold specification of the GORL indicated an increased likelihood of “severe injury” relative to “injury” outcomes.

4.1.4.2. Traffic characteristics

All three models suggest that, on average, lower speed limits have lower risk propensity relative to higher speed limits. To be specific, the negative coefficients of speed limits of 40 mph or less were found to be associated with lower likelihood of sustaining “injury” and “severe injury” relative to “no injury” outcomes according to the MNL framework. Similarly, both OR models had negative coefficients in their propensity equations indicating the lower risk towards higher severity outcomes as compared to the base case of 45 to 60 mph. It was not surprising that the involvement of a large truck in a work zone crash while traveling at higher speeds essentially proposed a deadly combination. Speed limits of 65 mph or higher, on the other hand, indicated the higher likelihood of higher severity outcomes relative to the base case category explained by the positive coefficients of all three models. This variable was found to influence injury between the “injury” and “severe injury” outcomes through the threshold parameter according to the GORL model.

TABLE 4 MNL, ORL, GORL model results

	MNL (Base Category: No Injury)		ORL	GORL	
Variable	**Injury**	**Severe Injury**	**Latent Propensity**	**Latent Propensity**	**Threshold: injury | severe injury**
Constant	-1.511 (-7.88)	-2.566 (-10.10)			
Roadway					
Functional class (base: collector, local system, rural minor arterial)					
Rural principal arterial	-	0.566 (2.81)	0.651 (3.22)	0.645 (3.21)	-
Urban principal arterial	-	-	0.454 (2.33)	0.434 (2.25)	-
Urban minor arterial	-	-	0.242 (1.29)	0.232 (1.24)	-
Geometric design					
Alignment (base: straight)					
Curved	-0.196 (-1.25)	0.283 (1.41)	-	-	-
Number of lanes (base: multi-lane)					
Two-lane	-	-0.445 (-1.95)	-	-	-
Curbed (base: no curb)					
Curb	-	-0.374 (-1.91)	-0.100 (-1.00)	-	0.214 (2.05)
Access-control (base: full control, and partial control)					
No control	0.263 (1.87)	0.950 (4.61)	0.654 (4.65)	0.612 (4.35)	-0.246 (2.42)
Traffic					
Speed limit (mph) (base: speed limit 45 to 60 mph)					
< 35 mph	-0.990 (-5.36)	-0.602 (-2.32)	-0.725 (-4.24)	-0.755 (-4.63)	-
35 - 40 mph	-0.366 (-1.61)	-0.651 (-1.78)	-0.398 (-1.93)	-0.404 (-1.98)	-
65 - 70 mph	-	0.579 (2.53)	0.284 (1.77)	0.232 (1.45)	-0.391 (-2.32)

TABLE 4 Continued

	MNL (Base Category: No Injury)		ORL	GORL	
Variable	**Injury**	**Severe Injury**	**Latent Propensity**	**Latent Propensity**	**Threshold: injury | severe injury**
Environmental					
Roadway surface condition (base: dry)					
Wet	-	-0.784(-2.25)	-0.369 (-2.03)	-0.348 (-1.92)	
Weather condition (base: clear)					
Adverse (rain, snow, fog, etc.)	-	0.455 (1.11)	0.230 (1.02)	0.213 (1.00)	-
Temporal					
Peak hours (base: off-peak)					
Peak	-0.187 (-1.63)	-	-0.162 (-1.63)	-0.156 (-1.57)	-
Time of the day (base: late night 12:00 AM - 6:00 AM)					
Day (6:00 AM - 6:00 PM)	0.400 (1.48)	0.972 (3.25)	0.567 (3.18)	0.493 (2.75)	-0.425 (-2.24)
Evening (6:00 PM - 12:00 AM)	0.277 (1.58)	-0.270 (-1.17)	-	-	-
Work Zone					
Workers (base: not present)					
Present	-	0.413 (2.60)	-	-	-
Work zone type (base: ln shift/crossover, intermittent/moving work zone)					
Lane closure	-0.279 (-2.42)	-0.349 (-2.05)	-0.236 (-2.21)	-0.245 (-2.30)	-
Shoulder or median	-	-	0.143 (1.20)	0.131 (1.11)	-
Work zone location (base: advanced-warning, activity, termination, other areas)					
Transition area	-0.238 (-1.66)	-0.615 (-2.70)	-0.373 (-2.94)	-0.375 (-2.97)	-
Crash					
Number of vehicles (base: multi-vehicle)					
Single-vehicle	-0.372 (-1.98)	-	-0.151 (-1.01)	-0.191 (-1.27)	-0.470 (-2.80)
Truck type (base: bus, 2 axels 1 unit, other)					
3+ axle 1 unit truck	-	0.384 (1.77)	-0.160 (-1.14)	-	-
1 unit Truck with trailer	-	0.468 (1.70)	0.354 (2.05)	0.335 (1.94)	-
Truck tractor semitrailer	-0.132 (-1.21)	-	-	-	-

TABLE 4 Continued

	MNL (Base Category: No Injury)		ORL	GORL	
Variable	**Injury**	**Severe Injury**	**Latent Propensity**	**Latent Propensity**	**Threshold: injury | severe injury**
Location					
Signalized intersection (base: no signal)	-	-0.172 (-1.06)	-	-	-
On-bridge (base: not on-bridge)	0.347 (1.78)	-	0.234 (1.34)	0.233 (1.33)	-
Threshold coefficients (ORL, GORL)					
No Injury | Possible Injury			0.4883 (3.78)	0.4660 (3.56)	
Possible Injury | Severe Injury			0.3134 (6.64)	0.4863 (6.61)	
Log-Likelihood at zero	-1915.10		-1,915.10	-1,915.10	
Log-Likelihood at convergence	-1836.62		-1,862.50	-1,847.37	
BIC	3,912.21		3,876.35	3,870.00	
McFadden R^2	0.0410		0.0275	0.0354	
Number of observations	2881		2881	2881	

A negative coefficient in the GORL threshold specification suggests lower likelihood of "injury" and higher likelihood of "severe injury" outcomes. Such a behavior was presented in earlier work zone crash severity literature; however a large truck involvement was not a factor (Li and Bai, 2009; Wang et al., 2010).

4.1.4.3. Environmental characteristics

Crashes on "wet surface" were associated with lower likelihood of "severe injury" relative to "no injury" in the MNL model. Similar results were obtained from the OR models. The ORL and GORL models indicate that roadways with wet surface have lower risk propensity relative to dry surface roadways. It seems as if truck drivers are more cautious driving at lower speeds and maintaining safe headways when driving on wet surface; such behavior has been suggested by past research (Chen and Chen, 2011; Duncan et al., 1998; Lemp et al., 2011; Zhu and Srinivasan, 2011a, 2011b).

Crashes during "adverse weather" conditions were associated with higher likelihood of sustaining "severe injury" relative to "no injury" according to the variable positive coefficient in the MNL model. The "adverse weather" variable was also found to be statistically significant with similar results obtained in the OR models. The ORL and GORL indicate that "adverse weather" has higher risk propensity relative to clear weather conditions indicated by the positive coefficients in their risk propensity functions. This result is consistent with earlier large truck crash severity literature; however these studies did not control for crashes specifically in work zones (Chang and Mannering, 1999; Chen and Chen, 2011; Dong et al., 2015; Wang and Shi, 2013). Adverse weather is likely to be associated with poor sight distance and visibility.

4.1.4.4. Temporal characteristics

Travelling during "peak-hours" was found to be associated with lower likelihood of "injury" relative to "no injury" according to the MNL. Similar results were obtained in the OR framework. To be specific, the negative coefficients of the ORL and GORL models indicate

lower risk propensity for traveling during peak-hours relative to non-peak hours. This is not a surprising result as traveling during peak-hours is typically congested leading to lower speeds, therefore reducing forceful impacts; such a result is consistent with past literature (Chang and Chien, 2013; Chang and Mannering, 1999; Duncan et al., 1998; M. Islam and Hernandez, 2013; Pahukula et al., 2015).

All three models showed that crashes during daytime were more severe compared to other times of the day. In the MNL, the magnitude of the positive coefficients indicated the higher likelihood of "severe injury" relative to "injury" outcomes. The negative coefficient in the GORL threshold specification essentially showed similar results. The "evening" indictor, in the MNL, was associated with higher likelihood of sustaining "injury" but lower likelihood of "severe injury" relative to "no injury" outcomes in the event of a crash. It seems that traveling at night can lead to an injury crash but not severe enough to cause severe injuries. Past studies have found similar results (Islam and Hernandez, 2013). Crashes during evening times are likely associated with lower visibility and higher speeds due to lower traffic volumes.

4.1.4.5. Work zone characteristics

The presence of worker in a work zone was associated with higher likelihood of "severe injury" relative to "no injury" outcomes according to the MNL model. The MNL positive coefficient value for the "severe injury" outcome essentially indicated that the presence of workers led to higher risks as they represent distraction to motorists.

Closing a lane or more in a work zone was found to be associated with lower likelihood of sustaining higher severity levels according to all three models. While the GORL failed to explain the effects of "lane closure" between the "injury" and "severe injury", the magnitude of the coefficients of both outcomes in the MNL indicated the lower likelihood of "severe injury" relative to "injury". Closing a lane or more is likely associated with the reduction of speed due to the combined traffic volumes into the functional lanes in a work zone.

Work on shoulders or medians led to higher severity levels in the event of a crash indicated by the positive coefficients in both of the risk propensity equations of the OR models. This higher risk is likely associated with travelling adjacent to fully functional lanes where large trucks tend to drive at higher speeds compared to partially or fully closed lanes.

Crashes in the transition area of a work zone were less likely to be severe as indicated by the negative coefficients in all three models. Drivers in the transition area have already passed through various advanced-warning and speed limit signs; the areas of a work zone that generally require lane changes and lane shifts, therefore motorist are likely to be already at lower speeds in those areas.

4.1.4.6. Crash characteristics

Crashes involving "single-vehicle" were found to be less severe according to the MNL and the risk propensity functions of both OR models; such a behavior was also suggested by earlier research (Qi et al., 2013). Interestingly, this variable had opposite effects in the GORL threshold equation between "injury" and "severe injury". Such a behavior suggests that although less involved vehicles can lead to lower likelihood of severe crashes, yet if an injury in fact occurred, the likelihood of "severe injury" is higher. Truck drivers are probably driving at higher speeds especially when not crowded by other vehicles in a work zone; therefore a sudden maneuver to change lanes or avoid workers could explain the opposite effects of the variable towards the lower and higher severity outcomes.

Crashes involving one-unit large trucks with three or more axles were found to be associated with higher likelihood of "severe injury" relative to the "no injury" outcomes indicated by the positive MNL coefficient for this variable. Past research have found similar result (Chen and Chen, 2011; Lemp et al., 2011). The more axles on a one-unit truck generally indicate heavier gross weight leading to forcible impacts. With a lower t-value in the ORL propensity equation, the negative coefficient value indicated the lower odds of higher severity levels; this behavior was

also found in the literature (Chang and Chien, 2013; Dong et al., 2015; Khorashadi et al., 2005; Zhu and Srinivasan, 2011a). Given the occurrence of a crash; the MNL relative to the ORL models suggest that if an individual in fact has sustained an injury, it is severe. Crashes involving one-unit trucks with trailers were more severe compared to other types of large trucks as indicated by the positive coefficients in all three models. Specifically, the MNL indicated the higher likelihood of "severe injury" relative to "no injury" outcomes as compared to buses, two-axle one-unit, and "other" types of large trucks. A trailer holds heavier cargo leading to higher severity levels. This result is consistent with the earlier research (Lemp et al., 2011; Zhu and Srinivasan, 2011a). The indicator of truck-tractor with a semitrailer was found to be associated with lower likelihood of "injury" relative to "no injury" according to the MNL negative coefficient; however, this result was associated with a lower statistical significance level. Previous large truck severity research suggested similar results; however these studies did not control for crashes specifically in work zones (Chen and Chen, 2011; Dong et al., 2015).

Signalized intersections were found to be associated with lower likelihood of "severe injury" relative to "no injury" outcomes as compared to non-signalized intersections according to the MNL model. Such a behavior was suggested by other studies; however these studies did not control for crashes specifically in work zones (Pahukula et al., 2015; Zhu and Srinivasan, 2011b).

All three models had positive coefficients for the "on-bridge" variable which essentially showed that crashes occurring on a bridge in a work zone were more severe compared to crashes on non-bridged roadways. Generally, bridges are poor locations for a large truck to maneuver, especially in a work zone where lane, shoulder, and median widths are usually kept at a minimum. This results is consistent with past work zone crash severity literature; however large truck involvement was not controlled for in those studies (Qi et al., 2013).

4.1.5. In-Depth Analysis of Large Truck Exposure over Time

The dataset utilized in this study is comprised of 10 years of large truck crashes in work zone. It is unknown to the authors how large truck exposure has changed over time between the beginning and the ending years of the study. It was necessary to further expand the GORL model to better capture the true effects of time on the severity of the most injured person in a crash. Table 5 presents further time-of-day analysis conducted within the GORL model. Based on the

TABLE 5 GORL-time-of-day interactions model results

	GORL	
Variable	**Latent Propensity**	**Threshold: injury | severe injury**
Roadway		
Functional class		
Principal arterial	0.609 (3.37)	-
Urban minor arterial	0.306 (1.66)	-
Geometric design		
Curbed		
Curb	-	0.253 (2.38)
Access-control		
No control	0.706 (5.31)	0.239 (2.31)
Traffic		
Speed limit (mph)		
< 45 mph	-0.663 (-4.54)	-
65 - 70 mph	0.267 (1.72)	-0.342 (-2.00)
Environmental		
Roadway surface condition		
Wet	-0.353 (-1.94)	0.133 (1.00)
Weather condition		
Adverse (rain, snow, fog, etc.)	0.230 (1.02)	-
Temporal		
Time-of-day		
Day (6:00 AM – 5:59 PM)	0.531 (3.02)	-0.514 (-2.52)
Year		
2003	0.297 (2.09)	-
2005	-	0.215 (1.36)
2006	-	1.046 (2.62)
2009	0.355 (2.24)	-
2010	0.314 (2.13)	-
Time-of-day and Year interactions		
Evening (6:00 PM – 11:59 PM) (2006)	-	-0.771 (-1.79)

TABLE 5 Continued

	GORL	
Variable	**Latent Propensity**	**Threshold: injury | severe injury**
Work Zone		
Work zone type		
Lane closure	-0.286 (-2.88)	-
Work zone location		
Transition area	-0.411 (-3.25)	-
Crash		
Number of vehicles		
Single-vehicle	-0.181 (-1.21)	-0.477 (-2.83)
Truck type		
1 Unit truck with trailer	0.230 (1.73)	-
Location		
On-bridge	0.245 (1.40)	-
Threshold coefficients		
No Injury | Possible Injury	0.6189 (5.84)	
Possible Injury | Severe Injury	0.3978 (4.78)	
Log-Likelihood at zero	-1,915.10	
Log-Likelihood at convergence	-1,839.13	
McFadden R^2	0.0397	
Number of observations	2,881	

hourly distribution of crashes within the dataset, several different categorizations of "time-of-day" variable were tested and compared using a Bayesian Information Criterion (BIC) test, discussed in a later section, in order to arrive at the best-fit distribution of crashes over the different times of the day. The overall effect of the different years within the dataset was also tested to investigate how the severity of crashes involving large truck within a work zone might have changed over time.

Finally, a partially-segmented GORL model was developed using interactions between "time-of-day" variable and "year" index in order to better address the effects of time layered within years in a composite way. The modified GORL model had positive coefficients for the years of 2003, 2009, and 2010 which essentially showed that crashes occurring during those years were more severe compared to crashes in other years within the dataset. On the other hand, years 2005 and 2006 indictors were associated with higher likelihood of sustaining "injury" but lower

likelihood of "severe injury" relative to "no injury" outcomes in the event of a crash. Those results do not indicate sufficient evidence that work zone enforcement practices have changed to the better or worse over the years of the current study. Interactions of "time-of-day" variable and "year" index showed statistical significance for evening crashes in the year of 2006 and late night crashes in the year of 2008 in which both variables essentially showed that crashes occurring during those specific time periods in both years were more severe compared to other time periods. Based on the results of the modified GORL model, an overall conclusion of interactions of "time-of-day" variable with "year" index is that truck exposure did not change during the different times of the day across the years in this study.

4.1.6. Measures of Fit

The MNL and ORL models cannot be compared using the log-likelihood ratio test statistic because they are non-nested models. Also, when fitting a set of models, it is possible to increase the goodness-of-fit by adding more parameters but this may result in obtaining an over-fitted model. The Bayesian Information Criterion (BIC) controls for over-fitting in a model by introducing a penalty term in its calculation, which essentially grows with adding more parameters to the estimated model (Akaike, 1987; Schwarz, 1978). The model with the lowest BIC value is essentially the best-fit among all. As shown in the model comparison table, the MNL, ORL, and GORL had BIC values of 3912.21, 3876.35, and 3870.00 respectively indicating that GORL has the lowest BIC value and thus provides superior data fit among the three models for modeling crash severity data of work zones involving large trucks.

4.1.7. Elasticity Effects

The magnitude of the effects of the independent variables entering a statistical model on each severity outcome is not directly provided through the parameter values provided by the model. To be able to clearly understand the impacts of these variables, it is necessary to compute

their corresponding elasticity effects. Elasticity effects can be interpreted as the percent effect a 1% change in a variable has on the severity outcome probability (Khorashadi et al., 2005). Elasticity calculations are not applicable to indicator variables; therefore average direct pseudo-elasticity was calculated (Chang and Mannering, 1999; Shankar and Mannering, 1996; Ulfarsson and Mannering, 2004). The pseudo-elasticity of a variable essentially represents the average percent change in the probability of an outcome category when the value of that variable is changed from 0 to 1. The elasticity analysis was undertaken only for the best model, *i.e.*, the GORL model.

4.1.7.1. Elasticity effects of GORL model

Aggregate level pseudo-elasticity effects of all the variables entered the GORL model were calculated and the results are shown in Table 6. The numbers in the top row of Table 6 indicate that the elasticity effects of "Rural principal arterial" functional class for "No Injury", "Injury", and "Severe Injury" outcomes are –15.14%, 50.35%, and 76.93%, respectively. So, work zone crashes involving large trucks occurring on rural principal arterials are 15.14% less likely to result in "no injury" whereas 50.35% and 76.93% more likely to result in "Injury" and "Severe Injury" outcomes respectively compared to crashes on collectors, local system roads, and rural minor arterial. Other numbers in the table can be interpreted similarly.

Based on the elasticity effects, it can be seen that the key factors and conditions that increase the risk of severe outcomes of crashes involving large trucks in work zones are: daytime crashes, no control of access, higher speed limit, and rural principal arterials. Other variables such as urban principal arterial, one-unit truck with trailer, and single-vehicle also contribute to increased risk, but not as much as the variables identified earlier.

Variable effects have important implications for training and education for drivers, workers, and non-motorists. These implications could also be extended to the planning and design of a work zone area and the regulation and use of traffic control devices. In terms of training and

TABLE 6 Elasticity effects of GORL

Variable	No Injury	Injury	Severe Injury
	Mean	Mean	Mean
Roadway			
Functional class (base = other=collector, local system, rural minor arterial)			
Rural principal arterial	-15.14	50.35	76.93
Urban principal arterial	-8.61	34.72	47.19
Urban minor arterial	-5.11	16.67	22.90
Geometric design			
Curbed (base=no curb)			
Curb	0.00	11.96	-23.38
Access-control (base=full control, and partial control)			
No control	-12.87	29.83	128.40
Traffic			
Speed limit (mph) (base=speed limit 45-60 mph)			
< 35	14.69	-42.91	-50.08
35 - 40	7.87	-25.39	-30.84
65 - 70	-5.15	-8.37	83.46
Environmental			
Roadway surface condition (base=dry)			
Wet	7.03	-22.03	-27.10
Weather condition (base=clear)			
Adverse (rainy, snowy, foggy, etc.)	-4.73	15.29	20.95
Temporal			
Peak hours (base=off-peak)			
Peak	3.32	-10.27	-13.13
Time of the day (base=late night 12:00 AM - 6:00 AM)			
Day (6:00 AM - 6:00 PM)	-11.58	3.77	134.94
Work Zone			
Work zone type (base=ln shift/crossover, intermittent/moving work zone)			
Lane closure	5.23	-15.74	-19.84
Shoulder or median	-2.84	9.37	12.49
Work zone location (base=advanced signs, activity, termination, other areas)			
Transition	7.59	-23.57	-28.87
Crash			
Number of vehicles (base=multi-vehicle)			
Single-vehicle	3.92	-34.18	35.69
Truck type (base=bus, 2 axels 1 unit, other)			
1 unit with trailer	-7.62	24.50	34.52
Location			
On-bridge (base=not on-bridge)	-5.21	16.79	23.08

education, the results suggest the importance of education to the drivers and training for work zone workers on daytime crash-developing situation in a work zone. It also suggests enforcing the use of highly reflective gears in work zones which increases the visibility of workers to the motorist.

In terms of planning and design, the results suggest that roadways with no control of access require assigning additional traffic control devices. It is well known in the transportation field that traffic control devices in work zones mandated by the FHWA are the minimum to be used; therefore extra traffic control measures may be warranted especially in areas with substantial large-truck traffic. Adding additional advisory and warning signs for non-motorists could effectively improve their alertness toward crash-developing situations. Speed harmonization methods and increased presence of law enforcement officers are recommended for enforcing lower speeds especially on non-controlled access roadways where more conflict points are present. Rerouting truck-traffic away from work zones on rural principal arterials could decrease the severity of a crash on this type of a functional class. Splitting truck traffic from other traffic will reduce conflicts in a work zone as well as give more space to non-heavy truck traffic for more flexible maneuvering to avoid possible crash situations. In terms of regulation of traffic, the results suggest to extend lower speed limits prior to entering work zone areas, which will allow more time for drivers to recognize the setup of the specific work zone being approached. It is essential to post traffic control signs that can communicate to vehicle drivers and non-motorists of sharing the roadway with large-truck traffic.

4.1.8. Conclusions

Safety literature focusing on work zone safety of large trucks is sparse. This research effort aims to fill this gap in the literature by undertaking an extensive empirical analysis of large truck crashes in work zones by pooling together 10 years of crash databases in the State of Minnesota. The empirical analysis employs statistical models that encompass recent advances in

the econometric literature. Specifically, both unordered and ordered modeling methods were deployed and the best modeling method for the current empirical context was chosen. To our knowledge, this is first such comparison of a comprehensive set of discrete choice models in the context of work zone safety.

A wide array of explanatory variables characterizing the crash, roadway, and work zone conditions were considered in the model estimation process. All models were gradually fine-tuned by removing statistically insignificant variables until the best-fit specification was obtained. In the unordered framework, the multinomial logit (MNL) and nested logit (NL) models were estimated. The NL model was used to test the validity of the IIA assumption in MNL model given the intrinsic ordinal nature of injury severity data being modeled. In the ordered response framework, simple ordered response logit (ORL) and generalized ordered response logit (GORL) models that explicitly recognize the ordinal nature of severity outcomes were estimated. The GORL model is a generalized version of the standard ORL model that relaxes the fixed thresholds assumption of the ORL thus providing additional flexibility. The performance of different models developed in this study was compared using Bayesian Information Criterion (BIC) test statistic. Among all the different models estimated in this study, the GORL model was found to offer the best-fit as indicated by its lower BIC value compared to other models. Lastly, going beyond simple parameter estimates, elasticity effects were computed to quantify the magnitude of impact of different exogenous factors considered in the study.

There are important empirical findings in the current study. The GORL model elasticity effects indicate that the most important factors/conditions that contribute to higher severity outcomes in the event of a crash are: daytime crashes, no control of access, higher speed limits, and crashes on rural principal arterials. Other variables such as urban principal arterial, one-unit truck with trailer, and single-vehicle also contribute to higher risk , but not as much as the variables identified earlier. With regards to potential improvements to this study, the authors used 10 years of crash data from the State of MN due to work zone data availability. So, the study

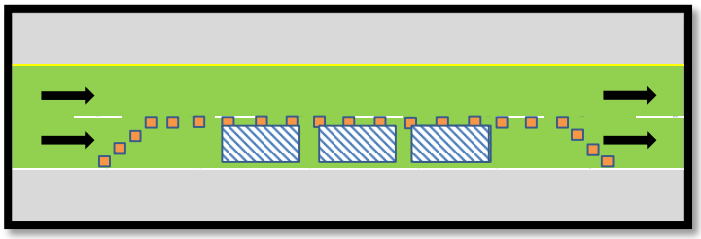
findings may not be extended to all work zones in the nation given that unique conditions specific to locations in MN may have influenced the analysis. Future research studies using combined datasets across multiple states will provide more evidence and confidence in the study findings. Also, bigger datasets allow segmentation of single and multi-vehicle crashes (*i.e.,* single truck crashes versus truck and car collisions) to check if there are significant differences in factors affecting severity of these two types of crashes. Another avenue for future research is exploring the endogeneity of work zone by including both work and non-work zone crashes in the analysis. Simultaneous modeling methods that jointly analyze crash occurrence at a work zone and severity conditional on crash occurrence in a work zone will enable unbiased estimation of model parameters (Eluru and Bhat, 2007; Kim and Washington, 2006). Future research including work zone-specific data such as modified lane, shoulder, and median widths, lengths of areas composing a work zone, and specific work zone speed limits could be beneficial. Also, in this study, we focused only on crash severity defined as the severity level of the most severely injured person in the crash. However, future studies must conduct occupant-level analysis that considers all people involved in the crash. This is important to obtain better insights into the relative profile of different occupant risk propensities and their determinants.

4.2. Injury Severity of passenger-Car Crashes for Different Work Zone Configurations

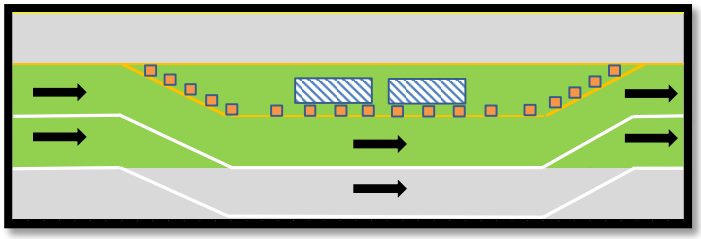
4.2.1. Introduction

Work zone safety remains a priority to the Federal Highway Administration, State Highway Departments, highway engineers, and the traveling public. Work zones create a hospitable environment for crashes; an issue that gained tremendous share of attention in recent years. Therefore, every effort should be sought out to reduce the injury severity of crashes in work zones. In this research we attempt to investigate the causal factors contributing to the injury severity of passenger-car crashes in different work zone configurations. A scan on the relevant literature indicates a gap in terms of analyzing the different work zone configurations encountered nationwide. The 2009 edition of the Manual on Uniform Traffic Control Devices (MUTCD) presents several applications of temporary traffic control (TTC) plans to address the different work zone needs based on the work being undertaken on roadways. Those TTC application are mandated nationwide and are considered the minimum any given work zone should apply ("Manual on Uniform Traffic Control Devices (MUTCD)," 2009). The majority of those TTC plans can generally be classified as one of the following five different work zone configurations: (1) lane closure, (2) lane shift/crossover, (3) shoulder/median, (4) intermittent/mobile, and (5) other. Fig. 2 demonstrates generic versions of each of the work zone configurations considered in this study with the exception of the "Other" category. Understanding the latent risks imposed by the specific configuration of the work zone, as a possible crash location, when interacting with other crash causal factors on the injury severity of crashes will allow practitioners, work zone designers, DOTs, to implement specific TTC devices, in addition to those recommended by the MUTCD which ultimately can mitigate those risks and therefore reduce the injury severity of involved occupants.

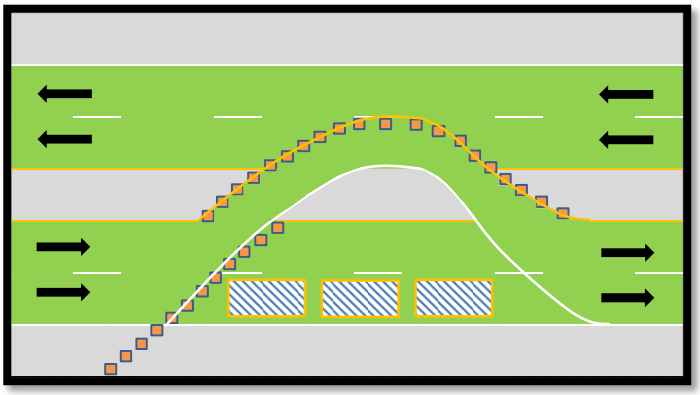
Considering the discrete ordinal nature of injury severity categories, a partially segmented Mixed Generalized Ordered Response Probit (MGORP) modeling framework was



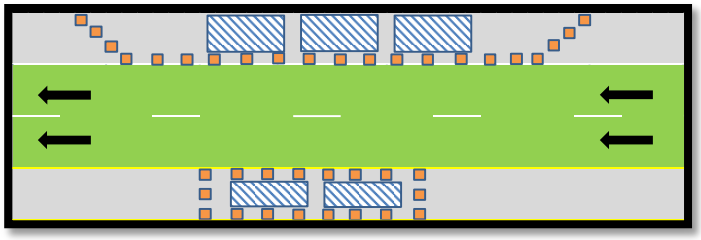
a. Lane Closure



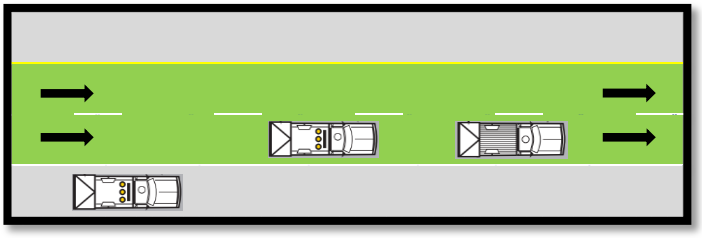
b. Lane Shift



c. Crossover



d. Shoulder or Median



e. Intermittent/Mobile

Fig. 2. Work zone configurations

(Adopted from: MUTCD 2009)

developed. The model estimation was undertaken by compiling a database consisting of 10 years of crashes that (1) involved at least one passenger car and (2) occurred in a work zone. Revealing the underlying factors contributing to injury severity levels pertaining to the most common work zone configurations will allow for the mitigation of higher severity outcomes.

4.2.2. Data

A dataset consisting of 10 years of work zone crashes (2003-2012) in Minnesota (MN) was collected from the HSIS database. The dataset contained 17,237 unique crashes reported to have occurred in work zones. Large-truck crashes was omitted from this study for three reasons: (1) low frequency of large-truck crashes, especially when interacted with the different work zone configurations, (2) approximately 85% of fatalities in work zones were drivers or occupants of passenger cars (FHWA, 2016), and (3) factors influencing the level of injury severity of involved individuals may vary significantly among truck versus non-truck crashes (Chang and Mannering, 1999). Although, truck involvement was accounted for as a binary variable in the modeling process in order to investigate whether it is in fact a risk factor contributing to the injury severity levels of occupants of passenger-cars in those cases involving both types of vehicles. The final sample of crashes was adjusted to 14,351 unique passenger-car crashes in work zones within the time frame depicted in this study. The distribution of observations by injury severity is presented in Tables 7 and 8. Table 7 shows the percentage of each injury severity category of the original dataset. The injury severity level followed the KABCO injury severity scale where K=killed, A=incapacitating injury, B=non-incapacitating injury, C=possible injury, and O=no injury. Due to the low frequency of some of the severity levels, some of the severity categories were combined. The combined injury severity categories are also shown in Table 8 Fatal, incapacitating, and non-incapacitating severity levels were combined into one severity level called “severe injury”. “Possible injury” which is referred to as “injury” and “no injury” categories were kept as is.

TABLE 7 Initial frequency of dependent variable (passenger-car)

Injury Severity Category	Count	(%)
Fatal (K)	63	0.44
Incapacitating Injury (A)	127	0.88
Non-Incapacitating Injury (B)	1,099	7.66
Possible Injury (C)	3,021	21.05
Property Damage (O)	10,041	69.97
Total	14,351	100.00%

TABLE 8 Final frequency of dependent variable (passenger-car)

Combined Injury Severity Category	Count	(%)
Severe Injury (K,A,B)	1,289	8.98
Mild Injury (C)	3,021	21.05
No Injury (O)	10,041	69.97
Total	14,351	100.00%

4.2.3. Empirical Analysis

Table 9 indicates the frequency distribution of the explanatory variables entered the MGORP modeling process. The authors adopted a methodological approach of interacting statistically significant factors with each of the five depicted work zone configurations, based on the specific work zone layout reported by the law enforcement agency investigating the crash. Differential impacts of the independent variables on the severity level were examined and the final specification for the presented model was based on a logical process of building a generalized ordered response probit (GORP) model while removing the statistically insignificant variables and combining other variables when their effects were statistically insignificant. Due to the complex process of crash occurrences to include, but certainly not limited to, interactions of vehicles, roadway conditions, traffic factors, and environmental conditions, it is considered almost impossible to gain access to all of the data contributing to the occurrence of a crash or its corresponding severity level. The lack of such important data can lead to erroneous specifications through biased parameter estimates (Mannering et al., 2016). This problem is typically referred to as “unobserved heterogeneity” in the crash analysis literature. We extensively tested for

TABLE 9 Frequency distribution of explanatory variable (passenger-car)

Explanatory Variable	(%)	Explanatory Variable	(%)
Roadway		**Work Zone**	
Geometric design		Work zone area	
Access control		Advanced-warning	11.20
No control	36.14	Transition	21.18
Partial control	9.33	Activity	64.05
Full control	54.53	Termination	3.57
Inclination		Work zone type	
On grade	24.15	Lane closure	38.90
Level	75.85	Lane shift/crossover	21.86
Alignment		Shoulder or Median	23.46
Curved	17.89	Intermittent	6.92
Straight	82.11	Other	8.86
No. of lanes		Presence of workers	
Two-lane	14.38	Workers present	32.28
Multi-lane	85.62	Workers not present	67.72
Roadway classification		**Temporal**	
Functional class		Day of the week	
Principal arterial	75.40	Weekday	79.23
Minor arterial	18.26	Weekend	20.77
Other (collector, local systems)	6.34	Time of day	
Area type		Daytime	73.39
Urban	85.14	Evening	19.93
Rural	14.86	Late night	6.68
Environmental		**Crash**	
Weather condition		No. of vehicles	
Adverse	35.84	Single-vehicle	21.73
Clear	64.16	Multi-vehicle	78.27
Roadway surface condition		Truck involvement	
Wet	18.83	Heavy-duty	3.87
Dry	81.17	Light-duty	33.56
Traffic		None	62.57
Speed limit		Location	
< 35	14.13	On-bridge	6.56
35-40	9.13	Not on-bridge	93.44
45-50	17.93		
55-60	52.02		
65-70	6.79		

unobserved heterogeneity effects of the injury severity determinants on the latent injury risk propensity due to potential unobserved factors. Thus, our final model specification became a partially segmented mixed generalized ordered response probit (MGORP) model. The final model

estimation process was, in large part, guided by findings of past research and intuitiveness of the parameters estimated. It terms of investigating the potential effects imposed by the specific work zone configuration where a crash has occurred, we followed a systematic approach of interacting all statistically significant variables with each of the five work zone configurations depicted in this study.

Fig. 3 represents the frequency of crashes within each of the five work zone configurations in the dataset. In the initial modeling process, each independent variable was regressed as a "standalone" variable to test for the statistical significance of its effect across all work zone configurations, followed by its additional interaction effects across each individual work zone configuration. The "other" work zone configuration served as the base for the remaining four categories for modeling specification purposes. For example, if a standalone variable had a coefficient parameter of +0.50 across all work zone configurations and its interaction with the "lane closure" configuration had an additional coefficient parameter of +0.15, the combined value of the two parameters (0.50 + 0.15 = +0.65) is the final effect of "lane closure" on this variable. Similarly, if the interaction of the same "standalone" variable with "shoulder or median" had a coefficient parameter of -0.20, therefore the combined effect for "shoulder or median" would be (0.50 - 0.20 = +0.30). This example can be interpreted as the "standalone" variable increased the likelihood of higher injury severity levels across all work zone configurations in the dataset with its positive coefficient value (+0.50). Relative to the "other" work zone configuration as the base category and compared to other work zone configurations, "lane closure" also increased the likelihood of higher injury severity with its positive coefficient (+0.15).

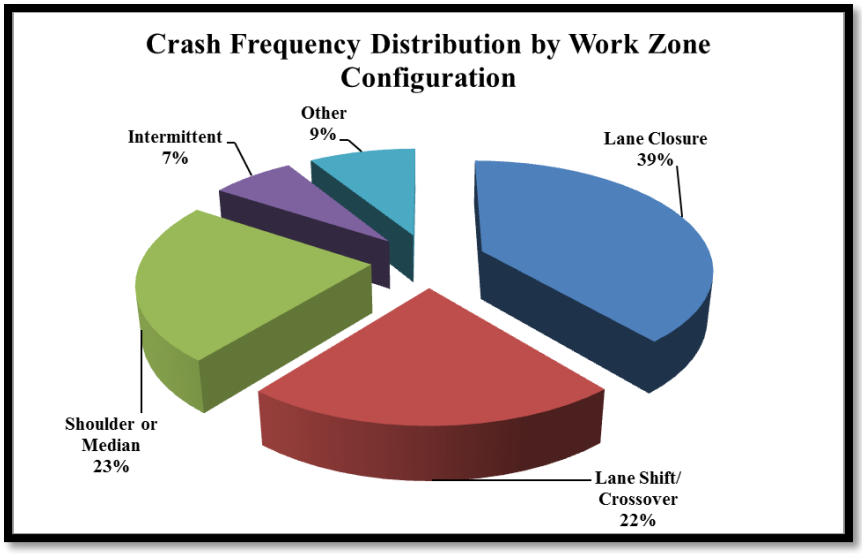


Fig. 3. Crash frequency distribution by work zone configuration

While "shoulder or median" also increased that likelihood with its positive coefficient (+0.30), it decreased the likelihood of higher injury severity levels relative to "other" work zone configurations with its negative interaction coefficient value (-0.20). This partially segmented approach uncovers the differences imposed by the different work zone configurations on each of the variables initially found statistically significant in the model before the introduction of any variable interactions.

4.2.4. Estimation Results

Table 10 presents the estimation results of the MGORP model. The first column of Table 10 shows the name of each variable entered the estimation process, while the second and third columns present two sets of variable coefficient parameters corresponding to the different injury severity levels. The second column of Table 10 presents each variable in the latent risk propensity function (excluding a constant) comparing the "no injury" vs. "injury" and "severe injury" outcomes. The third column of Table 10 presents variables entered the threshold specification function between "injury" and "severe injury" outcomes. Positive (+) parameter values indicate larger region of "injury" vs. "severe injury" under an injury severity curve, while negative (-)

parameter values indicate larger “severe injury” vs. “injury” outcomes. The respective t-values of the estimated coefficients are shown in parentheses. Table 10 also presents the initial log-likelihood value, the log-likelihood value at convergence, the McFadden R2, and the total number of observations in the dataset.

In the “variable” column, each variable is followed by its potential interactions with each of the different work zone configurations depicted in this study. For modeling specification reasons, the “other” category is considered the base for the remaining four work zone configurations throughout the modeling process. In the first column of Table 10, the four work zone configurations are demarcated by the numbers 1 thought 4 at the end of each variable’s name; lane closure (1), lane shift/crossover (2), shoulder/median (3), and intermittent/mobile (4).

4.2.4.1. Roadway characteristics

Relative to access-control “full control”, the positive parameters of “no control” and “partial control” indicated the increased risk propensity of higher injury severity outcomes. The negative threshold for “partial control” further indicated the increased proportion of “severe injury” relative to “injury” outcomes. Roadways with no access-control are likely to have more conflict points. While some studies indicated that full-control of access may contribute to the frequency of crashes in work zone (Khattak et al., 2002), there has not been any studies found in the work zone safety literature to address the accessibility of a roadway from an injury severity standpoint. Interactions of the “no control” variable with the different work zone configurations indicated that crashes occurred in lane closures were more severe while intermittent/mobile operations were associated with less injury severity relative to other work zone configurations. Lane closures in work zones with full-access to the roadway are likely to be associated with higher vehicular density in lanes open to traffic. Intermittent work zones in fully-accessed roadways are likely to be associated with lower vehicular speeds which can reduce forceful impacts at conflict points.

TABLE 10 MGORP model results (passenger-car)

	MGORP	
Explanatory Variables	**Latent Propensity**	**Threshold: injury | severe injury**
Roadway		
Geometric design		
Access control (base: full control)		
No control	0.241 (6.92)	-
No control-1	-	-0.083 (-1.38)
No control-4	-0.149 (-2.04)	
Partial control	0.175 (4.28)	-0.088 (-1.53)
Inclination (base: level)		
On grade	-	-0.057 (-1.42)
Alignment (base: straight)		
Curved	-	-0.136 (-2.86)
No. of lanes (base: multi-lane)		
Two-lane	0.117 (2.56)	-
Two-lane-1	-0.109 (-1.47)	-
Roadway classification		
Functional class (base: collector, local system)		
Principal arterial	0.070 (1.12)	-
Principal arterial-3	0.236 (3.80)	-
Minor arterial	0.218 (3.07)	-
Standard Deviation	0.343 (2.13)	
Minor arterial-1	-0.118 (-1.87)	-
Collector/local system-3	0.395 (3.36)	-
Area type (base: urban)		
Rural	-	-0.269 (-5.69)
Environmental		
Weather condition (base: clear)		
Adverse weather	-	0.111 (2.97)
Roadway surface condition (base: dry)		
Wet	-0.225 (-4.07)	-
Standard Deviation	0.314 (1.76)	
Wet-3	0.081 (1.20)	-
Traffic		
Speed limit (mph) (base: 45-60)		
< 35	-0.324 (-5.09)	-
Standard Deviation	0.592 (4.81)	
< 35-2	0.136 (1.45)	-
35-40	-	0.092 (1.69)
45-50-2	0.135 (2.28)	-
45-50-3	-0.243 (-3.48)	-
55-60-3	-0.185 (-2.91)	-
65-70	0.063 (1.38)	-

Interaction variables ending in 1-4 (1=lane closure, 2=lane shift/crossover, 3=shoulder/median, 4=intermittent/mobile)

TABLE 10 Continued

	MGORP	
Explanatory Variables	**Latent Propensity**	**Threshold: injury | severe injury**
Work Zone		
Work zone area (base: transition)		
Advanced-warning	0.205 (4.51)	0.065 (1.28)
Advanced-warning-3	-0.172 (-1.75)	
Activity	0.067 (2.10)	
Termination	0.115 (1.48)	-0.130 (-1.46)
Termination-3	0.160 (1.13)	
Termination-4	0.311 (1.57)	
Work zone type (base: shoulder/median, intermittent, other)		
Lane closure	-	0.073 (1.65)
Lane shift/crossover	-	0.090 (2.01)
Presence of workers (base: not present)		
Present	0.074 (2.28)	
Present-1	-0.050 (-1.10)	
Temporal		
Day of the week (base: weekday)		
Weekend	0.152 (5.43)	
Time of day (base: daytime)		
Evening	0.088 (2.96)	-0.104 (-2.46)
Late night	0.068 (1.28)	-0.287 (-4.06)
Late night-3	0.153 (1.42)	-
Crash		
No. of vehicles (base: multi-vehicle)		
Single-vehicle	0.069 (2.00)	-
Single-vehicle-1	0.183 (3.14)	-
Multi-vehicle-2	-0.077 (-1.97)	-
Truck involvement (base: none, light-duty)		
Heavy-duty	0.537 (6.93)	-0.213 (-2.59)
Heavy-duty-2	0.149 (1.10)	-
Heavy-duty-3	-0.308 (-2.26)	-
Location		
On-bridge (base: not on-bridge)		0.079 (1.21)
Constants		
Threshold 1 (no injury | injury)	-0.146	
Threshold 2 (injury | severe injury)	-0.074	
Log-Likelihood at	-11,399.9	
Log-Likelihood at convergence	-11,070.9	
McFadden R^2	0.0289	
Number of observations	14,351	

Interaction variables ending in 1-4 (1=lane closure, 2=lane shift/crossover, 3=shoulder/median, 4=intermittent/mobile)

For all work zone configurations, roadways on-grade and curved segments, as compared to "level" and "straight" respectively, increased the likelihood of higher injury severity outcomes (negative parameter values in the threshold function for both variables). Although some studies indicated the both curved and on-grade roadways increased the likelihood of single vehicle crash occurrences in work zones (Harb et al., 2008), yet there has been no comparative evidence to the findings of this study for the injury severity of crashes on both roadway alignments. Drivers are likely to be more cautious on a grade or a curved roadway, yet an unanticipated crash can lead to severe outcomes.

The number of lanes variable indicated that crashes occurring on two-lane roadways were associated with higher risk propensity of injury severity compared to multi-lanes roads. This finding is consistent with past literature for work zone crashes (Li and Bai, 2009). Interactions of the number of lanes with the different work zone configurations indicated that although crashes in the lane closure configuration were still associated with higher injury severities, yet the negative propensity specified that compared to other work zone configurations, lane closures reduced the severity of crashes. This is likely due to that fact that lane closures on a two-lane road is usually controlled with a temporary signal at the beginning and end of the work zone area so that one direction of traffic is traveling at a time across the work zone reducing conflicts with oncoming traffic. Past studies that accounted for "shoulder/median" activity in work zones did not provide conclusive evidence of contradicting results to the findings of this study, although rear-end crashes were controlled for (Qi et al., 2013).

Principal and minor arterials indicated an increased risk propensity towards higher injury severity outcomes compared to collectors and local systems. Previous studies (Li and Bai, 2008b; Qi et al., 2013) found similar results, which could be explained by higher speeds in the upstream area of a work zone. The standard deviation for the "minor arterial" variable indicated the presence of unobserved heterogeneity during the modeling process. Compared to all depicted

work zone configurations, shoulder/median activity on a principal arterial indicated higher risk towards higher injury severity levels. Although lane closures on a minor arterial still contributed to higher injury severity levels, its negative propensity indicated the reduced risk compared to other work zone configurations. Minor arterials are likely to have lower speed limits and higher vehicular density in work zones relative to principal arterials. Shoulder and median work on collectors or local system was associated with higher injury severity levels compared to other work zone configurations. This could be explained by the reduced availability of areas to maneuver (lack of shoulder or median) in a crash developing situation while likely traveling at maximum allowable speeds through the work zone in fully functional lanes adjacent to work areas.

Crashes occurring on roadways classified as "rural" indicated that in the event of a crash, the likelihood of the "severe injury" vs. "injury" outcomes is much higher. This was indicated by the negative coefficient parameter between both outcomes in the threshold function. This is likely due to higher speeds leading to a work zone area compared to an urban roadway. This finding is consistent with past work zone injury severity literature (Li and Bai, 2009; Qi et al., 2013; Wang et al., 2010) and work zone crash frequency literature (Khattak et al., 2002).

4.2.4.2. Environmental characteristics

Adverse weather and wet surfaces were associated with lower likelihood of severe injury crashes compared to clear weather conditions. It seems as if drivers are more cautious driving at lower speeds and maintaining safe headways when driving on wet surfaces or in an adverse weather situation. Other work zone studies found that wet surface had no impact on the severity of a crash relative to non-work zone areas (Harb et al., 2008; Li and Bai, 2009). Another study has found opposing results for fatal and injury crashes in work zones (Li and Bai, 2008b). Although traveling on wet surfaces in a work zone involving work on shoulder or median reduced

the risk of severe crashes, it appeared to be associated with least risk among other work zone configurations.

4.2.4.3. Traffic characteristics

Lower speeds upstream of work zones reduced risk propensities of higher severity crashes. This is indicated by the negative propensity of speeds under 35 mph. The positive coefficient in the threshold column for 35-40 mph indicated that if a crash occurred at those speeds, the most injured occupant would likely to sustain an injury but not a severe injury. When tested for unobserved heterogeneity, the standard deviation of 30 mph or less indicated strong statistical significance. On the other hand, positive coefficients in the risk propensity function for speed limits of 60 mph or more indicated a higher risk of a severe injury crashes. Previous work zone crash severity literature found similar results (Li and Bai, 2009; Wang et al., 2010). Interactions of the different speed limit categories with the different work zone configurations indicated that speeds of 50 mph or less were associated with higher severity outcomes in the event of a crash in lane shifts or crossovers compared to other work zone configurations. Lane shifts or crossovers are considered to be more complex work zone configurations relative to other types and are likely to be associated with potential distraction with machinery and workers ahead in the driver's line of sight. Speed limit range of 45-60 mph reduced the risk propensity of higher injury severity outcomes through work zones involving activity in the shoulder or median. Work outside the travel lane (i.e. shoulder or median) when balanced with mid-range speeds can lead to more attentive driving while allowing time and distance to come to a stop in a crash developing situation.

4.2.4.4. Work Zone characteristics

Advanced-warning, activity, and termination areas of a work zone were all associated with higher injury severity crashes indicated by the positive risk propensity coefficient values for all three variables as compared to the transition area. Motorists in the transition area are likely to

have already lowered their speeds after passing through advanced signage leading to the upcoming work zone and therefore, forceful impacts are reduced in the event of a crash. The activity area is likely to be associated with driver's distraction with work zone equipment and the presence of workers, while the termination area is likely to be associated with higher speeds exiting the work zone. No comparative evidence was found in the work zone injury severity literature to support or contradict such findings. One previous study concluded that the activity area was more susceptible to crashes regardless of the road type while the termination area had the lowest frequency (Garber and Zhao, 2002). Work on shoulder or median was associated with the least risk propensity of severe crashes in the activity area compared to other work zone configurations, yet increased the risk of those occurred in the termination area. This can be explained by the fact that drivers are likely to reduce their speeds approaching the advanced waning area of a work zone and easily gain speeds in the termination area especially when realized that actual work zone activity is not in the traveled lanes. The termination area of an intermittent/mobile operation was associated with higher risk propensity for higher severity outcomes compared to lane closures or lane shifts. Motorists are likely to encounter large pieces of moving equipment especially when merging into reopened lanes at the end of a moving work zone.

Compared to work on shoulder or median, intermittent/mobile, and "other" work zone configurations, lane closures and lane shifts were associated with injuries but not severe ones in the event of a passenger-car crash. This behavior was indicated by the model through the positive coefficient values of both variables in the threshold function. Although this study did not investigate the injury severity of work zone workers, it did account for their presence during the occurrence of a passenger-car crash in the work zone due to potential distraction to the driver. The presence of workers was associated with higher risk propensity of higher injury severity outcomes for passenger-car occupants. Presence of workers in a lane closure had the lowest risk among all other work zone configurations. The presence of workers generally represents a

distraction to drivers. Driver are likely to pay more attention to the specific location of workers to avoid striking one in case of possible intrusion of workers into the traveled lanes, and therefore less attention is given to other surroundings such as vehicles and traffic control devices.

4.2.4.5. Temporal characteristics

The "Weekend" indicator was found to be associated with higher likelihood of higher injury severity across all work zone configurations relative to traveling on the weekdays. Past literature indicated similar results for non-truck involved crashes (Chang and Mannering, 1999). Such a behavior is likely due to the fact that most, and certainly not all, work zones are inactive during weekends. Motorists would likely speed through the work zone once discovered it is not operational. Higher speeds will lead to forceful impacts in the event of a crash.

Traveling during evening and late night times increased the propensity risks of severe passenger-crashes in work zones compared to daytime crashes. The positive risk propensities and negative thresholds values, in the MGORP model results indicated such a behavior. The highly significant negative threshold value for "late night" indicated that in the event of a crash, a passenger-car occupant is likely to sustain "severe injury" relative to "injury" outcomes. Past work zone safety studies found similar results for night time crashes (Chang and Mannering, 1999; Harb et al., 2008). This can be explained by poor visibility at late night times and higher speeds due to lower vehicular densities compared to daytime. Although most work zones are inactive during late night times, the work zone configuration of shoulder or median work was associated with the highest risks among other configurations. Shoulder or median work zone configurations involve the least exposure to work zone objects (e.g. cones, barriers, attenuators) in the traveled lanes and therefore motorists are likely to raise speeds due to less intimidation by conflicts.

4.2.4.6. Crash characteristics

The "single-vehicle" crash indictor was found to be associated with higher risk propensities for passenger-car crashes in work zones. Single-vehicle crashes in lane closures were associated higher risks compared to other work zone configurations. Single-vehicle crashes usually involve inattentive driving and in the case of lane closures, sudden maneuvers to change lanes or avoid equipment or worker's intrusion in the travel lane are expected. No comparative evidence in the work zone injury severity literature to support or contradict such findings. Drivers are probably riding at lower speeds especially when crowded by other merging vehicles in a lane shift configuration.

Compared to light-duty trucks and other passenger-cars involved in work zone crashes, the involvement of heavy-duty trucks was found to be highly associated with higher risk propensities. The highly significant value of the "heavy-duty" variable in the risk propensity function indicated such a behavior. The negative threshold value for the same variable further indicated that in the event of a passenger-car crash involving a heavy-duty truck, passenger-car occupants are likely to sustain "severe injury" rather "injury" outcomes. Such a behavior was suggested by past literature (Chang and Mannering, 1999; Harb et al., 2008; Li and Bai, 2009), while another study suggested opposing results, although some of these studies did not control for crashes specifically in work zones (Chen and Chen, 2011; Dong et al., 2015; Wang et al., 2010). Being fatigued or falling asleep is not unusual among truck drivers (Saltzman and Belzer, 2007), although these conditions are not particular to just work zones. Interactions of the "heavy-duty" indicator with the different work zone configurations indicated that lane shifts or crossovers were associated with higher risk propensities of severe crashes while work on shoulder/median reduced that risk. Heavy-duty trucks are likely harder to maneuver when shifting lanes compared to other work zone configurations, especially during a sudden reaction to another vehicle or workers in a work zone. Although the MGORP model failed to provide a coefficient in the risk propensity function for "on-bridge", the positive coefficient value of the threshold function indicated that in

the event of a passenger-car crash on a bridge, the outcome is an injury rather than a severe injury. Drivers are likely to lower their speeds crossing an active work zone on a bridge therefore; forceful impacts are unlikely to occur in the event of a crash.

4.2.3. Measures of Fit

Table 10 indicates that the MGORP has a log-likelihood (LL) value of -11,070.9 at convergence. Comparatively, the model with constants in the threshold function and no covariates in risk propensity has a LL value of -11,399.9. The MGORP model has 54 additional parameters compared to the constants only model. The Likelihood Ratio (LR) test statistic of comparison between the MGROP and the constants-only model was 658 which is greater than the chi-squared critical value of 72.15 (at the 0.05 level of significance) corresponding to 54 degrees of freedom.

4.2.4. Elasticity Effects

The magnitude of the effects of the independent variables entering a statistical model on each injury severity outcome is not directly provided through the parameter values produced by the model. To be able to clearly understand the impacts of these variables, some of which appear in both the risk propensity and the threshold functions for the MGORP model, it is necessary to compute their corresponding elasticity effects. Elasticity effects can be interpreted as the percent effect a 1% change in a variable has on the severity outcome probability (Khorashadi et al., 2005). Elasticity calculations are not applicable to indicator variables; therefore average direct pseudo-elasticity was calculated (Li and Bai, 2008a; Wong et al., 2011; Zhu and Srinivasan, 2011a). The pseudo-elasticity of a variable represents the average percent change in the probability of an outcome category when the value of that variable is changed from 0 to 1. The elasticity analysis was undertaken for the MGORP model and the results are shown in the following subsection.

4.2.4.1. Elasticity effects of MGORP model

Elasticity effects were calculated for all three injury severity outcomes. For the sake of brevity, only results corresponding to the "severe injury" outcome category are presented herein (see Table 11). The first five columns in Table 11 present the results in cases where the elasticity effects vary across different work zone configurations whereas the last column shows the elasticity effects for variables whose impacts do not have such a variation across the different work zone configurations.

In terms of elasticity effects of variables that do not vary across different work zone configurations, the first value is the last column of Table 11 corresponding to "partial control" is 51.20. This indicates that occupants of passenger-car crashes are 51.20% more likely to be severely injured in the event of a crash occurring in work zones in roadways with access-control "partial control" relative to "full control". Moreover, this effect does not vary across different work zone configurations. In Table 11, elasticities of other variables that do not vary across the different work zone configurations in the last column can be interpreted in a similar fashion. Elasticity effects of variables that vary across work zone configurations indicates, the involvement of heavy-duty trucks is found to impose the highest risk of severe outcomes in "lane shift/crossover" (186.39%) followed by other configurations, while "shoulder/median" has the least risk (47.56%) among all other work zone configurations. In Table 11, elasticities of other variables that varied across different work zone configurations in the first five columns can be interpreted similarly.

Based on the elasticity effects in Table 11 , it can be seen that the key factors and conditions that increase the risk of severe outcomes for the occupants of passenger-cars across all work zones are: partial control of access, roadways classified as rural, crashes during evening times, crashes during weekends, and curved roadways. Other variables such as crashes in the activity area of a work zone, higher speeds of 65-70 mph, and roadways on a grade also contribute to increased risk, but not as much as the variables identified earlier. Variations in

elasticity effects of variables across different work zone configurations were found for the following factors – access-control, number of lanes, roadway functional class, roadway surface condition, speed limit, work zone component area, presence of workers in the work zone, time-of-day, number of involved vehicles, and truck involvement.

TABLE 11 Elasticity effects of MGORP model for "Severe Injury" outcome (passenger-car)

Explanatory Variable	Lane Closure	Lane Shift / Crossover	Shoulder / Median	Intermittent	Other	Main Variable Effects
Roadway						
Geometric design						
Access control (base: full control)						
No control	70.07	51.90	51.90	17.88	51.90	
Partial control						51.20
Inclination (base: level)						
On grade						8.67
Alignment (base: straight)						
Curved						21.23
No. of lanes (base: multi-lane)						
Two-lane	1.43	22.21	22.21	22.21	22.21	
Roadway classification						
Functional class (base: other=collector, local system)						
Principal arterial	13.12	13.12	67.13	13.12	13.12	
Minor arterial	18.84	44.42	44.42	44.42	44.42	
Other	0.00	0.00	88.32	0.00	0.00	
Area type (base: urban)						
Rural						44.03
Environmental						
Weather condition (base: clear)						
Adverse weather						-15.65
Roadway surface condition (base: dry)						
Wet	-33.34	-33.34	-22.53	-33.34	-33.34	
Traffic						
Speed limit (mph) (base: 45-60)						
< 35	-44.46	-28.27	-44.46	-44.46	-44.46	
35-40						-13.35
45-50	0.00	25.43	-35.78	0.00	0.00	
55-60	0.00	0.00	-28.17	0.00	0.00	
65-70						11.47

TABLE 11 Continued

Explanatory Variable	Lane Closure	Lane Shift / Crossover	Shoulder / Median	Intermittent	Other	Main Variable Effects
Work Zone						
Work zone area (base: transition)						
Advanced-warning	40.99	40.99	5.96	40.99	40.99	
Transition	0.00	0.00	-19.01	44.35	0.00	
Activity						12.38
Termination	21.74	21.74	57.39	97.30	21.74	
Work zone type (base: shoulder/median, intermittent, other)						
Lane closure						-10.48
Lane shift/crossover						-12.93
Presence of workers (base: not present)						
Present	4.29	13.54	13.54	13.54	13.54	
Temporal						
Day of the week (base: weekday)						
Weekend						29.68
Time of day (base: daytime)						
Evening						34.73
Late night	12.49	12.49	44.74	12.49	12.49	
Crash						
No. of vehicles (base: multi-vehicle)						
Single-vehicle	52.70	12.88	12.88	12.88	12.88	
Multi-vehicle	0.00	-12.69	0.00	0.00	0.00	
Truck involvement (base: none, light-duty)						
Heavy-duty	134.64	186.39	47.56	134.64	134.64	
Location						
On-bridge (base: not on-bridge)						-11.63

4.2.5. Implications of Variable Effects and Recommendations

Variable effects have important implications for the regulation and use of traffic control devices based on the general configuration of the work zone it is, and generally for planning and design of work zones. These implications could also be extended to the training and education for drivers, work zone workers, and non-motorists. In the context of this research, these implications can be classified into two categories: (1) across all work zone configurations, and (2) across specific work zone configurations.

4.2.5.1. Recommendations for all work zone configurations

In terms of TTC regulation and use across all work zone configurations, the modeling results suggest that on roadways that lacks full control-of-access, additional TTC signage and warning messages are needed in the upstream areas of access-points to advice motorists of such upcoming locations. A speed limit reduction shall be mandated and enforced, and not just recommended, upstream of work zones on roadways classified as "rural". There shall be additional lighting enforcement practices during the evening times, especially if the work zone is active. The condition and reflectivity of TC devices shall be strictly maintained and the usage of additional warning lights to clearly demarcate travel lanes from work areas in the evening times is encouraged. Substantial consequences shall be executed by DOTs towards those who are found in violation. When feasible, means of Intelligent Transportation Systems (ITS) (e.g. digital message boards (DMSs)) shall be employed to communicate operations that are active during the evening times. Heftier fines shall be imposed on speeding motorists despite the fact that a work zone may or may not be operational during weekends. Although most work zones are not operational during weekends, it shall be clearly communicated to motorists if it is in fact active. Direct communication with motorists in the vicinity of a work zone (e.g. message boards, DMSs), or with potential off-site motorists (e.g. social media, radio stations) is encouraged which may divert such motorists from joining the work zone.

4.2.5.2. Recommendations for specific work zone configurations

In terms of planning and design, the results suggest that splitting heavy-duty truck traffic from other traffic will reduce conflicts, especially when lane shifts exist within a work zone. If at all feasible, this suggestion shall be extended to other work zone configurations. The results suggest that the transition and termination areas of the intermittent/mobile work zone configuration shall be extended beyond MUTCD recommended lengths; this is to allow clearance distance from any moving equipment so that motorists can make a safer merging maneuver out of an occupied lane or into reopened lanes past the work zone. Enforcing lower speeds in the termination area of the "shoulder or median" configuration shall be assured until the work zone is entirely crossed. This can be established through mandating the presence of law enforcement officers at the end of the work zone; typically all lanes are fully open to traffic in such a work zone configuration. Lower speed limits shall be posted and enforced beyond those recommended by the MUTCD, especially for work zones involving shoulder or median activities on collectors or location systems among all other work zone configurations. The existence of lane closures in roadways with no-access-control shall be clearly communicated to traffic joining the mainline at conflict points. This can be established through the usage of message boards in the upstream area of access points. This practice shall be extended to the "lane shift/crossover", "shoulder/median", and "other" work zone configurations which are also associated with increased risks for severe crashes but not as much as lane closures. As compared to all other work zone configurations, shoulder/median activities on principle arterials shall be clearly communicated to motorists in the advanced warning area. Work on shoulder or median are associated with fully functional travel lanes and therefore motorists are likely to speed through the work zone not knowing work exists, but not in the traveled lanes. Additionally, law enforcement presence on principal arterials upstream of the advanced-warning area for "shoulder or median" work activities is recommended for the enforcing of reduced speed limits. On minor arterials, similar practices shall be introduced in all work zone configurations but not necessarily in the "lane closure" configuration which is

associated with the least injury severity risks among all. Shoulder or median activities shall be clearly communicated to motorists traveling late at night, which are associated with the highest risk among all other work zone configurations. Work activities on shoulders or medians are not obvious to motorists traveling at late night times as much as other work zone configurations.

In terms of training and education, the results suggest the importance of education for motorists and training for the personnel of the agency overseeing the operation within the work zone. It is essential to install TTC devices that can communicate to the motorist the specific configuration of the work zone being approached. Work zone safety seminars shall be offered to the traveling public to teach them about what may be different once a work zone is erected in their community, and how this may affect their daily commute. FHWA mandates only minimum traffic control applications for different work zones, therefore additional traffic control devices and measures may be warranted especially for unique features of potential configurations. Training for government agency personnel or their representative (e.g. work zone safety classes) shall be mandated vs. recommended; this is crucial in terms of learning about the different factors affecting the severity of crashes within certain work zone configurations. Learning more about these factor will allow for the recognition of potential hazardous situations, and therefore the tailoring of additional counter measures pertaining to the specific work zone configuration in effect. For example, work zone managers should learn the additional TTC signage needed when more than anticipated heavy-duty truck traffic is present in the work zone.

4.2.6. Conclusions

While safety literature, to date, is considered fairly rich in analyzing many aspects of crashes in work zones, it certainly lacks attention-to-details in terms of risks imposed on the injury severity level of crashes occurring in different work zone configurations. This research effort aims to fill this gap in the literature by undertaking an extensive empirical analysis of passenger-car crashes in work zones by using 10 years of crash databases in the State of

Minnesota. The authors wish to investigate the most contributing factors affecting the injury severity level of the most injured occupant of passenger-cars involved in crashes in work zones. The empirical analysis employs the mixed generalized ordered response probit model (MGORP) model that recognizes the ordinal nature of the data while allowing for heterogeneity to capture the effects of unobserved factors. The primary focus of this study is to uncover the potential interaction effects that the different common work zone configurations impose on the factors contributing to the crash. In doing so, effects of regressed variables were taken into consideration while revealing additional effects produced through interactions to finally produce the net effect for each variable within each specific work zone configuration. In the context of work zone safety, and to author's knowledge, this is the first study to explore the factors affecting the injury severity of occupants of passenger-cars at the level of individual work zone configurations commonly recognized by the MUTCD.

The MGORP model that accounts for unobserved heterogeneity and threshold heterogeneity across crashes was found to fit the utilized dataset while addressing limitations imposed by simpler modeling techniques in past injury severity literature (i.e. ordered probit model (ORP)). Also, there are several important empirical findings in the current study. The MGORP model elasticity effects indicates that key factors that increases the likelihood of severe crashes includes – partial control of access, roadways classified as rural, crashes during evening times, crashes during weekends, and curved roadways. Other variables such as crashes in the activity area of a work zone, higher speeds of 65-70 mph, and roadways on a grade also contribute to higher risks, but not as much as the variables identified earlier. Although, these variables were common to all work zones.

With regards to variations across the different work zone configurations, significant differences were observed in the effects of the following factors – access-control, number of lanes, roadway functional class, roadway surface condition, speed limit, work zone component

area, presence of workers in the work zone, time-of-day, number of involved vehicles, and truck involvement.

One of the limitations of this study was that there were very few variables in the database describing the work zone-specific features (for example, work zone duration, lane and shoulder widths). In terms of future research, the collection of work zone-specific data such as work zone-specific lane, shoulder, and median widths, lengths of areas composing a work zone, work zone duration, and specific work zone speed limits could be beneficial to provide more insights to design ideal work zone parameters for enhancing traffic safety.

4.3. Injury Severity of drivers for Different Work Zone component-Areas

4.3.1. Introduction

According to the 2009 edition of the Manual on Uniform Traffic Control Devices (MUTCD), and as shown in Fig. 4, a work zone mainly consists of the following components : (1) advance warning area, (2) transition area, (3) activity area, and (4) termination area ("Manual on Uniform Traffic Control Devices (MUTCD)," 2009). Each of these areas has a specific purpose and may vary in size and location depending on the nature of the activity being done.

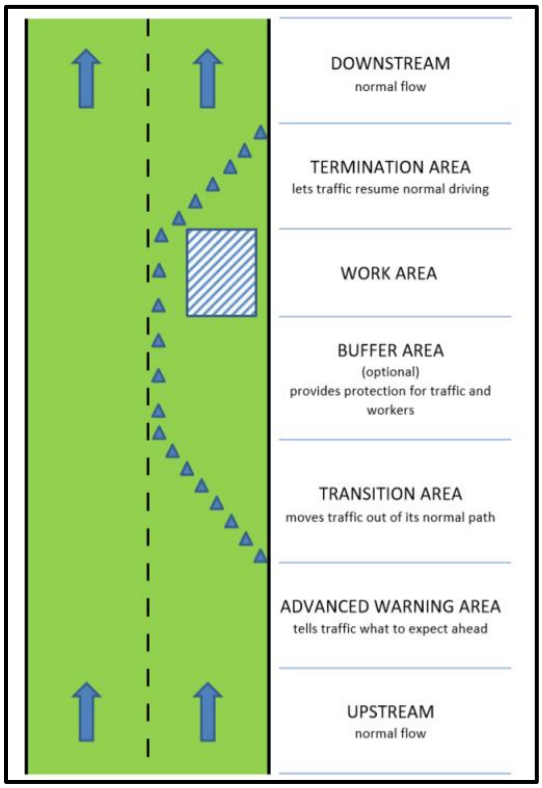


Fig. 4. Work zone component-areas
(Adopted from: MUTCD 2009)

According to current work zone safety literature to date, work zone specific component-areas were never the subject of analysis in the context of crash injury severity. Thus, the current study aims to contribute to the literature on work zone safety by exploring the characteristics of

crashes occurring in the different areas composing a work zones through using a disaggregate-level analytical approach. The severity of each individual driver involved in a work zone crash in the State of Minnesota (2003-2012) was investigated through a set of potentially associated factors that may contribute to the degree of severity in which each driver has sustained. Specifically, the study examines the factors that impact the driver's injury severity level when involved in a work zone crash and the potential additional impacts of this crash occurring in a specific work zone area. Additional effects are investigated through methods of interactions between potentially contributing factors and each or the work zone component –areas.

Understanding the different characteristics affecting the severity of driver in the different work zone areas will serve as a great advantage enabling practitioners, designers, and DOT officials to mitigate the injury severity level of those drivers generally involved in a work zone crash or particularly within a specific component-area of the work zone or both.

4.3.2. Data

A dataset consisting of 10 years of work zone crashes (2003-2012) in Minnesota (MN) was collected from the Highway Safety Information System (HSIS). The dataset contained 28,358 drivers involved in crashes occurred in work zones. The driver severity level followed the KABCO injury severity scale where K=killed, A=incapacitating injury, B=non-incapacitating injury, C=possible injury, and O=no injury. The distribution of observations by injury severity is presented in Tables 12 and 13. Table 12 shows the percentage of each severity category of the original data. Due to the low frequency of some of the severity levels, some of the severity categories were combined. The combined injury severity categories are also shown in Table 13 Fatal, incapacitating, and non-incapacitating severity levels were combined into one severity level called "severe injury". "Possible injury" which is referred to as "injury" and "no injury" categories were kept as is.

TABLE 12 Initial frequency of dependent variable (driver)

Injury Severity Category	Count	(%)
Fatal (K)	43	0.15
Incapacitating Injury (A)	94	0.33
Non-Incapacitating Injury (B)	1,019	3.60
Possible Injury (C)	3,160	11.14
Property Damage (O)	24,042	84.78
Total	28,358	100.00%

TABLE 13 Final frequency of dependent variable (driver)

Combined Injury Severity Category	Count	(%)
Severe Injury (K,A,B)	1,156	4.08
Mild Injury (C)	3,160	11.14
No Injury (O)	24,042	84.78
Total	28,358	100.00%

4.3.3. Empirical Analysis

Table 14 indicates the frequency distribution of the explanatory variables entered the MGORP modeling process. The authors adopted a methodological approach of interacting statistically significant factors with each of the four work zone areas, based on the exact work zone area each crash has occurred in. Differential impacts of the independent variables on the severity level were examined and the final specification for the presented model was based on a logical process of building a GORP model while removing the statistically insignificant variables and combining other variables when their effects were statistically insignificant. The authors extensively tested for unobserved heterogeneity effects of the injury severity determinants on the latent injury risk propensity due to potential unobserved factors. Thus, our final model specification became a partially segmented mixed generalized ordered response probit model (MGORP). The final model estimation process was, in large part, guided by findings of past research and intuitiveness of the parameters estimated.

TABLE 14 Frequency distribution of explanatory variable (driver)

Explanatory Variable	(%)	Explanatory Variable	(%)
Crash-level		**Temporal**	
Roadway		Day of the week	
Geometric design		Weekday	80.5
Roadway grade		Weekend	19.5
Grade	23.3	Time of day	
Level	76.7	Daytime	76.4
Roadway division		Evening	18.9
Divided	77.0	Late night	4.7
Undivided	23.0	**Crash**	
Access control		No. of vehicles	
No control	34.6	Single vehicle	11.0
Partial control	9.1	Multi vehicle	89.0
Full control	56.3	**Vehicle/occupant-level**	
Roadway classification		**Driver**	
Functional class		Residence	
Principal arterial	76.8	Out-of-state	18.0
Minor arterial	18.0	In-state	82.0
Other	5.2	Gender	
Area type		Male	41.1
Urban	86.6	Female	58.9
Rural	13.4	Age	
Environmental		Younger	35.9
Weather condition		Middle	49.7
Adverse	35.0	Older	14.4
Clear	65.0	Seatbelt usage	
Roadway surface condition		Used	90.5
Wet	17.0	Not used	9.5
Dry	83.0	Alcohol usage	
Traffic		Used	2.3
Speed limit (mph)		Not used	97.7
< 35	13.1	Ejected	
35-40	9.5	Yes	7.9
45-50	18.4	No	92.1
55-60	52.3	**Vehicle**	
65-70	6.7	Airbag deployment	
Work Zone		Deployed	8.5
Work zone type		Not deployed	91.5
Lane closure	40.5	Vehicle type	
Lane shift/crossover	20.8	Passenger car	60.9
Shoulder/median	23.8	Truck-light duty	36.1
Intermittent	6.7	Truck heavy duty	3.0
Other	8.2	Vehicle age (years)	
Work zone area		> 10 years	67.8
Advanced-warning	12.3	< 10 years	32.2
Transition	21.7	Number of occupants	
Activity	62.6	Single-occupant	74.8
Termination	3.4	Multi-occupant	25.2

Fig. 5 represents the frequency of crashes within each of the four work zone component-areas in the dataset. In the initial modeling process, each independent variable was regressed as a "standalone" variable to test for the statistical significance of its effect across all work zone component-areas combined, followed by its additional interaction effects across each individual work zone component-area. The "termination" work zone component-area served as the base for the remaining three categories for modeling specification purposes. For example, if a standalone variable had a coefficient parameter of +0.50 across all work zone component-areas combined and its interaction with the "activity" component-area had an additional coefficient parameter of +0.15, the combined value of the two parameters (0.50 + 0.15 = +0.65) is the final effect of "activity" on this variable.

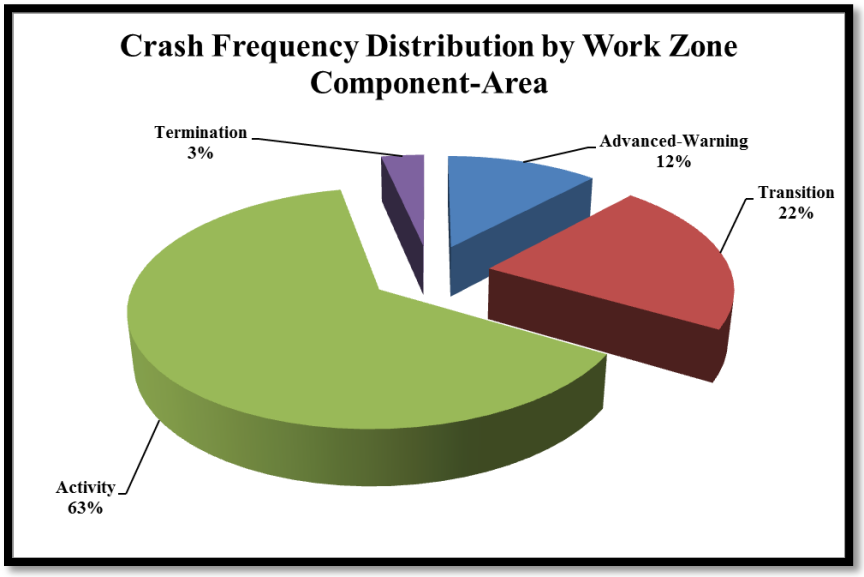


Fig. 5. Crash frequency distribution by work zone component-area

Similarly, if the interaction of the same "standalone" variable with "transition" had a coefficient parameter of -0.20, therefore the combined effect of "transition" on the "standalone" variable would be (0.50 - 0.20 = +0.30). This example can be interpreted as the "standalone" variable increased the likelihood of higher injury severity levels across all work zone component-areas in the dataset with its positive coefficient value (+0.50). Relative to the "termination" work

zone component-area as the base category and compared to other work zone areas, “activity” also increased the likelihood of higher injury severity with its positive coefficient (+0.15). While “transition” also increased that likelihood with its positive coefficient (+0.30), it decreased the likelihood of higher injury severity levels relative to the “termination” work zone component-area with its negative interaction coefficient value (-0.20). This partially segmented approach uncovers the differences imposed by the different work zone component-areas on each of the variables initially found statistically significant in the model before the introduction of any variable interactions.

4.3.4. Estimation Results

Table 15 presents the estimation results of the MGORP model. The first column of Table 15 shows variables’ names, while the second and third columns present two sets of variable coefficients corresponding to the different severity levels. The second column presents each variable in the latent risk propensity (excluding a constant) comparing the “no injury” vs. “injury” and “severe injury” outcomes. The third column present the variables entered the threshold specification function between “injury” and “severe injury” outcomes. The respective t-values of the estimated coefficients are shown in parentheses. Table 15 also presents the initial log-likelihood value, the log-likelihood value at convergence, the McFadden R^2, and the total number of observations in the dataset.

In the “variable” column, each variable is followed by its potential interactions with each of the areas comprising a work zone. The four work zone areas are demarcated by the numbers 1 thought 4 at the end of each variable’s name; advanced warning area (1), transition area (2), activity area (3), and termination area (4). The “termination area” or “area 4” is considered the base category for all interaction variables throughout the modeling process.

4.3.4.1. Roadway characteristics

Roadways on a grade, as compared to "level", decreased the likelihood of higher injury severity outcomes (negative coefficient value in the latent risk propensity function). Undivided roadways and roadways with partial- or no control-of-access increased the risk propensity of higher injury severity. The negative threshold coefficient for undivided roadways indicated that crashes were more severe relative to divided roadways. A median would reduce conflict points. Non-access-controlled roadways are likely to have more conflict points. While the interaction of no control-of-access with the advanced warning area still followed the same injury severity direction as the rest of the work zone areas, its negative coefficient in the propensity column indicated that drivers, in this work zone area, sustained less severe injuries than all other areas.

Principal and minor arterials indicated a higher driver's risk propensity towards higher injury severity outcomes levels relative to collectors and local systems. Previous studies (Li and Bai, 2008b; Qi et al., 2013) found similar results, which could be explained by higher speeds in the upstream area of a work zone. Urban roadways indicated less likelihood of higher injury severity outcomes with its negative risk propensity value. While urban roadways are likely to carry more congested traffic, speeds are typically lowered relative to rural areas in work zones.

4.3.4.2. Environmental characteristics

Adverse weather and wet surfaces were associated with lower likelihood of drivers sustaining higher injury severity compared to clear weather conditions. It seems as if drivers are more cautious driving at lower speeds and maintaining safe headways when driving on wet surfaces or in an adverse weather situation. Other work zone studies found that wet surface had no impact on the severity of a crash relative to non-work zone areas (Harb et al., 2008; Li and Bai, 2009). Another study has found opposing results for fatal and injury crashes in work zones (Li and Bai, 2008b).

4.3.4.3. Traffic characteristics

Lower speeds upstream of a work zones reduced driver's risk propensities. The positive coefficient in the threshold column for 35-40 mph indicated that if a crash occurred at those speeds, drivers would likely to sustain an injury but not a severe injury. On the other hand, negative coefficients in the threshold function for speed limits of 45 mph or more indicated a higher risk of a severely injured driver. These results were not surprising as the involvement in a crash in a work zone while being subject to interactions with heavy equipment. Previous work zone crash severity literature found similar results (Li and Bai, 2009; Wang et al., 2010).

4.3.4.4. Work zone characteristics

Lane closures in work zones was found to be associated with lower likelihood of sustaining higher injury severity levels according to the positive threshold value for this variable. Closing a lane or more is likely associated with the reduction of speed due to the combined traffic volumes into the functional lanes. Intermittent operations were found to be associated with drivers sustaining higher injury severity. Specifically, the negative value of the "intermittent" variable in the threshold function indicated a higher likelihood of "severe injury" relative to "injury". In an intermittent operation, drivers are likely to interact with additional traffic control devices in the work zone as compared to stationary operations areas; such as truck- mounted attenuators, flaggers, and message boards mounted on light or heavy duty trucks directing traffic and protecting workers in the activity area the operation moves ahead. Interactions between the "work zone type" variables and the four areas composing a work zone indicated that the "activity area" of a moving operation has the lowest risk on driver's injury severity among advanced, transition, and termination work zone areas. This is likely due to that fact that in the activity area of a moving operation, drivers have already passed through any needed lane changes and reached their lowest speed through decelerating in the advanced warning and the transition areas. Also, the activity areas in an intermittent operations work zone are likely to occupy shorter repair

TABLE 15 MGORP model results (driver)

	MGORP	
Variable	**Latent Propensity**	**Threshold: injury | severe injury**
Crash-level		
Roadway		
Geometric design		
Roadway grade (base: level)		
On grade	-0.048 (-1.65)	
Roadway division (base: divided)		
Undivided	0.066 (1.67)	-0.055 (-1.37)
Access control (base: full control)		
No control	0.279 (6.65)	
No control-1	-0.199 (-2.47)	
Partial control	0.157 (3.47)	
Roadway classification		
Functional class (base: other)		
Principal arterial	0.107 (1.70)	
Minor arterial	0.071 (1.20)	
Area type (base: rural)		
Urban	-0.123 (-3.33)	
Environmental		
Weather condition (base: clear)		
Adverse weather		0.091 (2.47)
Roadway surface condition (base: dry)		
Wet	-0.225 (-3.04)	
Standard Deviation	0.544 (3.67)	
Traffic		
Speed limit (mph) (base: More than 35)		
< 35	-0.300 (-6.40)	
35-40		0.108 (1.97)
45-50		-0.076 (-1.72)
65-70		-0.137 (-3.05)
Work Zone		
Work zone type (base: shoulder/median and other)		
Lane closure		0.049 (1.23)
Lane shift/crossover		0.120 (2.62)
Intermittent		-0.215 (-1.98)
Intermittent-3		0.184 (1.35)
Temporal		
Day of the week (base: weekend)		
Weekday	-0.209 (-2.78)	
Standard Deviation	0.553 (4.85)	
Weekend-1	-0.143 (-1.44)	
Weekend-3	-0.137 (-2.13)	
Time of day (base: evening, and late night)		
Daytime	-0.070 (-2.27)	0.073 (1.87)
Late night -1	-0.542 (-2.32)	

Interaction variables ending in 1-4 (1=advanced-warning, 2=transition, 3=activity, 4=termination (base))

TABLE 15 Continued

Variable	Latent Propensity	Threshold: injury | severe injury
Late night -2	0.192 (1.53)	
Crash		
No. of vehicles (base: single vehicle)		
Multi vehicle	-0.389 (-8.68)	0.059 (1.24)
Multi vehicle-1	0.141 (1.97)	
Occupant/vehicle-level		
Driver		
Residence (base: in-state)		
Out-of-state	-0.406 (-4.48)	
Standard Deviation	0.607 (4.10)	
Out-of-state-2	-0.152 (-1.67)	
Gender (base: female)		
Male	0.270 (8.84)	0.189 (5.27)
Female-1		0.286 (4.24)
Female-2	-0.115 (-2.50)	
Age (base: middle: 30-60 years old)		
Younger (< 30 years old)	-0.119 (-4.27)	
Older (> 60 years old)	-0.126 (-2.11)	-0.111 (-2.21)
Older (> 60 years old)-3	0.110 (1.54)	
Seatbelt usage (base: seatbelt used)		
Seatbelt not used		-0.181 (-2.81)
Alcohol usage (base: alcohol not used)		
Alcohol used	0.487 (6.70)	
Ejection (base: not ejected)		
Ejected	0.197 (4.02)	-0.252 (-3.72)
Vehicle		
Airbag (base: not deployed)		
Deployed	1.126 (16.54)	
Vehicle type (base: passenger car)		
Truck-light duty	-0.094 (-3.48)	
Truck-heavy duty	1.112 (8.94)	
Truck-heavy duty-3	-0.259 (-2.05)	
Vehicle age (years) (base: < 10 years)		
> 10 years	-0.202 (-2.91)	0.124 (2.24)
Standard Deviation	0.817 (6.78)	
No. of occupants (base: single-occupant)		
Multi-occupant	0.121 (3.80)	0.087 (2.22)
Single-occupant-1	0.167 (2.22)	
Constants		
Threshold 1 (no injury | injury)	-0.104	
Threshold 2 (injury | severe injury)	-0.246	
Log-Likelihood at zero	-14,827.8	
Log-Likelihood at convergence	-13,440.9	
McFadden R^2	0.0935	
Number of observations	28,358	

Interaction variables ending in 1-4 (1=advanced-warning, 2=transition, 3=activity, 4=termination (base))

segments compared to those of the stationary work zone type. Shorter-in-length activity areas lead to drivers spending less time through the work zone in a moving operation.

4.3.4.5. Temporal characteristics

"Weekdays" was found to be associated with lower likelihood of higher injury severity relative to traveling on the weekend according to the MGORP model results. "Weekday" variable showed statistical significance when tested for heterogeneity (S.D. = 0.553). Interactions of "weekend" variable indicated that the advanced-warning and the activity areas decreased driver's risk of high injury severity relative to other work zone areas. Such a behavior was likely due to the fact that most motorists would lower their speeds entering the advanced warning areas but once discovered that the work zone is not operational at the time, drivers are likely to speed through the transition area. The lower risk propensity associated with the activity area is likely due to fewer conflicts with workers and heavy construction equipment during downtime on weekends.

Traveling during "daytime" was found to be associated with lower likelihood of higher driver's injury severity in work zones (negative coefficient value in the risk propensity function). Similarly, the positive coefficient in the threshold function indicated that in the event of a crash, a driver would sustain an injury relative to a severe injury. Traveling during the day is likely associated with congested roadways; therefore, lower speeds would reduce forceful impacts. Interactions of the time-of-day variable indicated that the advanced warning areas had an increased likelihood of drivers sustaining higher injury severity while the transition area lowered this risk.

4.3.4.6. Crash characteristics

The "multi-vehicle" crash indictor was found to be associated with driver's lower risk propensities. Such a behavior was also suggested by earlier research (Qi et al., 2013). This was indicated by both the positive coefficient value of the risk propensity function as well as the

negative value of the threshold function. Drivers are probably riding at lower speeds especially when crowded by other vehicles in a work zone. In the event of a single-vehicle crash, sudden maneuvers to change lanes or avoid equipment or worker's intrusion in the travel lane are expected. Interaction of the advanced-warning area with "multi-vehicle" indicated the lowest driver's risk propensity relative to other work zone areas.

4.3.4.7. Driver attributes

"Out-of-state" indicator had lower risk propensities, although this variable showed statistical significance when tested for heterogeneity effects (S.D. 0.607). These results are consistent with previous literature (Harb et al., 2008). Interactions of the "out-of-state" variable indicated that the transition area had the lowest risk propensity relative to other work zone areas. An out-of-state driver is expected to be more cautious paying additional attention to traffic control devices due to unfamiliarity with the area. As the transition area starts after the advanced-warning area, and given that drivers are likely already driving at lower-than-the posted speed limits leading to a work zone, it is not surprising that the safest work zone area would be the transition area relative to interacting with the presence of worker and heavy equipment in the activity area, similarly speeding up to normal speeds through the termination areas.

Relative to female drivers, the male drivers had higher risk propensity (positive coefficient). Interestingly, the positive threshold value for the "male" variable has a monotonic effect which indicated that although males are riskier, that in the event of a crash, they would sustain just an injury relative to a severe injury. Physically, female drivers are susceptible to higher injury severities. Previous literature have found similar results (Weng and Meng, 2011). Gender interactions indicated that females have lower likelihood of sustaining severe injuries in both advanced warning and transition areas.

Relative to middle age drivers, both younger and older drivers indicated a lower risk propensity for sustaining higher injury severity. The negative value in threshold function for older

drivers indicated that this age group is likely to sustain “severe injury” relative to “injury” outcomes. Such results could be explained by the reduced risks taken by both age groups compared to the middle age group who would likely take higher risks driving at higher speeds through the work zone. These results are consistent with previous literature (Weng and Meng, 2011). Interactions of “older driver” indicated a higher likelihood of higher injury severity outcomes in the activity area relative to all other areas, which could be explained by being in close range with distractions such as heavy machinery in the activity area.

Lack of seat belt usage was found to be associated with higher driver’s injury severity. The negative coefficient value in the threshold function indicated that drivers not using their seat belt have higher likelihood of sustaining “severe injury” relative to “injury” outcomes. The indicator for “alcohol used” indicated that drivers had higher risk propensities when under the influence. Previous literature found similar results (Harb et al., 2008). Drivers ejected from a vehicle, in the event of a crash, had higher risk propensity for higher injury severity outcomes. The negative coefficient value in the threshold function for “ejected” indicated higher risks for severe injuries relative to just an injury. Ejection into a work zone would especially increase the chance of being impacted by machinery, or other devices.

4.3.4.8. Vehicle characteristics

Deployment of airbags increased driver’s risk propensity outcomes. The airbag “deployed” variable is a unique one in term of the way researchers would interpret it. It is well known that airbags are usually deployed as a result of an impact and not necessarily a contributor to the crash cause. In the context of this study, the authors interpreted airbag deployment as a sign of severe impacts. Previous literature confirmed an assumption that the deployment of airbags would reduce crash fatalities among belted drivers, yet the risk was increased among unbelted drivers (Høye, 2010).

Light-duty trucks reduced driver's risk propensity of higher injury severity. These results are consistent with the literature (Chang and Mannering, 1999). Heavy-duty trucks had higher risks (Chang and Mannering, 1999; Dong et al., 2015). Being fatigued or falling asleep is not unusual among truck drivers (Saltzman and Belzer, 2007), although these conditions are not particular to just work zones. Interactions of the "truck-heavy duty" revealed that the activity area had the lowest risk propensity among all other areas. Heavy-duty trucks are likely at their lowest speeds in the activity areas, which would reduce forceful impacts in the event of a crash.

Older vehicles were associated with lower severity levels among drivers. Additionally, the positive threshold coefficient value for vehicle age over 10 years indicated that in the event of a crash, a driver would likely to be injured but not severely. The severity of the driver's injury is likely associated with the vehicle's body and frame material composition. The automotive industry and manufacturers have been leaning towards using lighter-weight materials in newer vehicles for better benefits such as better fuel economy, drivability, and performance (Cole and Sherman, 1995). It is intuitive that a more solid built vehicle (i.e. steel or cast iron) would protect its driver from severe impacts relative to light-weight vehicles (i.e. aluminum and magnesium alloy) (Cole and Sherman, 1995; Miller et al., 2000). Some previous literature found opposite results (Weng and Meng, 2011). Vehicle age variable was found to be statistically significant when tested for heterogeneity (S.D. 0.817).

"Multi-occupant" vehicles were more severe compared to single-occupant. This behavior could be explained by the fact that additional persons in a vehicle represent a distraction to the driver. The positive threshold coefficient value for "multi-occupant" indicated that additional persons in a vehicle led to a driver's injury relative to severe injury outcomes. Additional passengers in a vehicle might warn the driver of an oncoming danger overlooked by the driver (i.e. another vehicle). Previous literature found similar results (Khattak and Targa, 2004). Interactions of "single-occupant" indicated that the advanced warning area had higher risks of severe outcomes compared to other work zone areas. Advanced-warning areas mainly consist of

open roadways with reflective signage indicting that a work zone is being approached. Inattentive drivers could miss signage leading to a work zone.

4.3.5. Measures of Fit

Table 15 indicates that the MGORP has a log-likelihood value of (-13,440.9) at convergence. Comparatively, the model with constants in threshold and no covariates in risk propensity has a LL value of -14,827.8. The MGORP model has 56 additional parameters compared to the constants only model. The Likelihood Ratio (LR) test statistic of comparison between the MGROP and the constants only model was 2773.8 which is greater than the chi-squared critical value of 74.47 (at the 0.05 level of significance) corresponding to 56 degrees of freedom. We also estimated the fixed parameters GORP model that ignores unobserved heterogeneity. The LL value of the GORP model was (-13,457.12) and the LR test statistic of comparison with MGORP model was 32.44. This value is greater than the critical chi squared value of 9.49 corresponding to 4 degrees of freedom. This demonstrates superior data fit in the MGORP model and the importance of accounting for unobserved heterogeneity in injury severity models.

4.3.6. Elasticity Effects

The magnitude of the effects of the independent variables entering a statistical model on each severity outcome is not directly provided through the parameter values in the model. To be able to clearly understand the impacts of these variables some of which appear in both the risk propensity and the threshold functions for the MGORP model, it is necessary to compute their corresponding elasticity effects. Elasticity effects can be interpreted as the percent effect a 1% change in a variable has on the severity outcome probability (Khorashadi et al., 2005). Elasticity calculations are not applicable to indicator variables; therefore average direct pseudo-elasticity was calculated (Li and Bai, 2008a; Wong et al., 2011; Zhu and Srinivasan, 2011a). The pseudo-elasticity of a variable essentially represents the average percent change in the probability of an

outcome category when the value of that variable is changed from 0 to 1. The elasticity analysis was undertaken for the MGORP model and the results are shown in the following subsection.

4.3.6.1. Elasticity effects of MGORP model

Elasticity effects were calculated for all three injury severity outcomes. For the sake of brevity, only results corresponding to the "severe injury" outcome category are presented in the paper (see Table 16). The first four columns in Table 16 present the results in cases where the elasticity effects vary across different work zone areas whereas the last column shows the elasticity effects for variables whose impact does not vary across different work zone areas.

The first value is the last column of Table 16 corresponding to 'on grade' roadway is -8.79. This indicates that drivers are 8.79% less likely to be severely injured in the event of a work zone crash occurring on "on-grade" relative to "level" roadways. Moreover, this effect does not vary across different work zone areas. Similarly, results corresponding to access control suggest that drivers involved in work zone crashes in "advanced-warning" area are 16.85% more likely to sustain severe injuries compared to crashes in work zone areas with full access control. Furthermore, the effect of access control also varies across different work zone areas. Specifically, drivers involved in crashes at work zones with no access control are 70.40% more likely to sustain severe injuries in "transition", "activity", and "termination" areas compared to 16.85% in "advanced-warning" area. Other values in Table 16 can be interpreted similarly.

Overall, significant variations in elasticity effects across work zone areas were found for the following factors – type of work zone, time of day, number of vehicles involved in the crash, gender, age, and residence status of the driver, and type of vehicle. Also, it can be seen that the key factors and conditions that increase the risk of severe outcomes of among drivers involved in

TABLE 16 Elasticity effects of MGORP model for "Severe Injury" outcome (driver)

Variable	Work Zone Area Interactions				Main Variable Effects
	Advanced Warning	Transition	Activity	Termination	
Crash-level					
Roadway					
Geometric design					
Roadway grade (base: level)					
On grade					-8.79
Roadway division (base: divided)					
Undivided					25.23
Access control (base: fully controlled)					
No control	16.85	70.40	70.40	70.40	
Partial control					34.30
Roadway classification					
Functional class (base: other)					
Principal arterial					22.68
Minor arterial					14.51
Area type (base: rural)					
Urban					-20.82
Environmental					
Weather condition (base: clear)					
Adverse weather					-15.57
Roadway surface condition (base: dry)					
Wet					-35.47
Traffic					
Speed limit (mph) (base: > 35)					
< 35					-44.31
35-40					-18.63
45-50					14.43
65-70					26.36

TABLE 16 Continued

Variable	Work Zone Area Interactions				Main Variable Effects
	Advanced Warning	**Transition**	**Activity**	**Termination**	
Work Zone					
Lane closure					-8.52
Lane shift/crossover					-20.30
Intermittent	42.33	42.33	5.82	42.33	
Temporal					
Day of the (base: weekend)					
Weekday					-32.69
Weekend	-24.02	-	-23.07	-	
Time of day (base: evening, late night)					
Daytime					-23.43
Late night	-66.86	43.16	-	-	
Crash					
No. of vehicles (base: single vehicle)					
Multi vehicle	-37.49	-52.88	-52.88	-52.88	
Occupant/vehicle-level					
Driver					
Residence (base: in-state driver)					
Out-of-state	-55.40	-67.67	-55.40	-55.40	
Gender (base: female)					
Male					18.19
Female	-44.61	-19.86	-	-	
Age (base: middle age: 30-60 years old)					
Younger (< 30 years old)					-20.38
Older (> 60 years old)	-21.61	-21.61	-2.97	-21.61	
Seatbelt usage (base: seatbelt used)					
Seatbelt not used					36.35
Alcohol (base: alcohol not used)					
Alcohol used					146.22

TABLE 16 Continued

Variable	**Work Zone Area Interactions**				**Main Variable Effects**
	Advanced Warning	**Transition**	**Activity**	**Termination**	
Ejection (base: not ejected)					
Ejected					120.97
Vehicle					
Airbag (base: not deployed)					
Deployed					772.77
Vehicle type (base: passenger car)					
Truck-light duty					-16.60
Truck-heavy duty	593.90	593.90	366.95	593.90	
Vehicle age (years) (base: < 10 years)					
> 10 years					-45.23
No. of occupants (base: single-occupant)					
Multi-occupant					7.12
Single-occupant	36.85	-	-	-	

work zones crashes are airbag deployment, alcohol involvement, driver ejection out of vehicle, lack of seatbelt usage, and partial access of the work zone.

4.3.7. Conclusions

Safety literature is sparse on econometric modeling of crashes occurring in different areas within a work zone. This research effort aims to fill this gap in the literature by undertaking an extensive empirical analysis of crashes in work zones by using 10 years of crash databases in the State of Minnesota. The authors wish to investigate the most contributing factors affecting the injury severity of drivers involved in crashes within a work zone. The empirical analysis employs the mixed generalized ordered response probit model (MGORP) model that recognizes the ordinal nature of the data while allowing for heterogeneity to capture the effects of unobserved factors. The primary focus of this study is to uncover the potential interaction effects that the nature of each work zone area imposes on the factors contributing to the crash. In doing so, effects of regressed variables were taken into consideration while revealing additional effects produced through interactions to finally produce the net effect for each variable within each specific work zone area. In the context of work zone safety, and to our knowledge, this is the first study to explore the factors affecting driver injury severity at the level of individual work zone areas.

The MGORP model that accounts for unobserved heterogeneity and threshold heterogeneity across crashes was found to fit the data significantly better than the standard ORP model. Also, there are several important empirical findings in the current study. The MGORP model elasticity effects indicates that high-impact crashes involving airbag deployment, alcohol, driver ejection, lack of seatbelt usage, and partial access control of work zone are key factors and conditions that increase the risk of severe outcomes among drivers involved in work zones crashes.

With regards to variation across work zone areas, significant differences were observed in the effects of the following factors – type of work zone, time of day, number of vehicles involved in the crash, gender, age, and residence status of the driver, and type of vehicle.

One of the limitations of this study was that there were very few variables in the database describing the work zone conditions (for example, work zone type is one such variable). In terms of future research, the collection of work zone-specific data such as work zone-specific lane, shoulder, and median widths, lengths of areas composing a work zone, and specific work zone speed limits could be beneficial to provide more insights to design ideal work zone parameters for enhancing traffic safety.

5. CONCLUSIONS AND FUTURE RESEARCH

5.1. Conclusions

This dissertation proposed a comprehensive discrete choice analysis of injury severity of crashes in work zones. Part of this dissertation focused on investigating several different econometric frameworks that had been previously utilized in the field of injury severity analysis while proposing newer and more robust modeling analysis that addressed most of the limitations found in previous modeling structures. Injury severity of individuals involved in work zone crashes was analyzed for occupants of large truck crashes, occupants of passenger-car crashes and finally for all drivers involved in both large trucks and passenger-car combined.

Safety literature focusing on work zone safety of large trucks is sparse. Part of this dissertation aimed to contribute to the literature in this field by conducting an extensive empirical analysis of large truck crashes in work zones by pooling together 10 years of crash databases in the State of Minnesota. A comparison between four different econometric frameworks was undertaken to distinguish the best-fit modeling structure for crash data analysis. Specifically, both unordered and ordered modeling methods were employed and the best modeling method for the current empirical context was chosen. This is first such comparison of a comprehensive set of discrete choice models in the context of work zone safety.

Based on the superiority of the GORL model in the large truck crash study, another part of this dissertation focused on analyzing the injury severity of occupants of passenger-cars in work zone crashes. In doing so, the GORL model was restructured to the MGORP model to further capture heterogeneity in the effects of covariates while providing the same robust features of the GORL among all other utilized models in the previous part of this dissertation. The MGORP was utilized to investigate the additional effects of different work zone configurations imposed on the covariates found significant in the study through means of interactions between those covariates and each of the work zone configurations. There has not been any studies found

in the literature of work zone safety that analyzed injury severity of work zone crashes for different work zone configurations.

Another part of this dissertation focused on analyzing the injury severity of drivers involved in work zone crashes on as occupant-level bases. The MGORP modeling structure was also used for this study. This study also focused on uncovering the additional effects of the different work zone component-areas imposed on the factors contributing to the injury severity of drivers involved in work zone crashes. Means of interactions between statistically significant variables and each of the different work zone component-areas were utilized to capture both the effects of occupant-level risk factors as well as any additional effects imposed by the specific work zone component-area where a crash has occurred.

Important empirical findings were identified and discussed for each of the three studies composing this dissertation. Various variable implications were identified for training and education of motorists, non-motorists, workers, government agencies overseeing work zone operations. In terms of planning an work zone design, other recommendations were made for work zone designers, and government agencies regarding additional traffic control measures the can mitigate the severity of crashes in work zones based on the risk factors influencing large trucks crashes and passenger cars in different work zone configurations.

5.2. Avenues for Future Research

Future research studies using combined datasets across multiple states will allow for extending the study findings nationwide given that unique conditions specific to locations in MN may have influenced the analysis. Also, bigger datasets allow segmentation of single and multi-vehicle crashes (*i.e.*, single truck crashes versus truck and car collisions) to check if there are significant differences in factors affecting severity of these two types of crashes.

Another avenue for future research is exploring the endogeneity of work zone by including both work and non-work zone crashes in the analysis. This is important because injury

severity outcomes at a work zone site can be more (or less) severe because of unobserved factors that caused the site to be a work zone site. Simultaneous modeling methods that jointly analyze crash occurrence at a work zone and severity conditional on crash occurrence in a work zone will enable unbiased estimation of model parameters (Eluru and Bhat, 2007; Kim and Washington, 2006).

Future research including work zone-specific data such as modified lane, shoulder, and median widths, lengths of areas composing a work zone, and specific work zone speed limits could be beneficial. Also, in this study, we focused only on crash severity defined as the severity level of the most severely injured person in the crash. However, future studies can conduct occupant-level analysis that considers all people involved in the crash. This is important to obtain better insights into the relative profile of different occupant risk propensities and their determinants.

REFERENCES

Akaike, H., 1987. Factor analysis and AIC. Psychometrika 52, 317–332. doi:10.1007/BF02294359

Akepati, S.R., Dissanayake, S., 2011. Risk Factors Associated with Injury Severity of Work Zone Crashes. Presented at the Transportation Research Board 90th Annual Meeting.

Anowar, S., Yasmin, S., Tay, R., 2014. Factors Influencing the Severity of Intersection Crashes in Bangladesh. Asian Transp. Stud. 3, 143–154. doi:10.11175/eastsats.3.143

Arditi, D., Lee, D.-E., Polat, G., 2007. Fatal accidents in nighttime vs. daytime highway construction work zones. J. Safety Res. 38, 399–405. doi:10.1016/j.jsr.2007.04.001

Bédard, M., Guyatt, G.H., Stones, M.J., Hirdes, J.P., 2002. The independent contribution of driver, crash, and vehicle characteristics to driver fatalities. Accid. Anal. Prev. 34, 717–727. doi:10.1016/S0001-4575(01)00072-0

Bhat, C.R., 2001. Quasi-random maximum simulated likelihood estimation of the mixed multinomial logit model. Transp. Res. Part B Methodol. 35, 677–693. doi:10.1016/S0191-2615(00)00014-X

Chang, L.-Y., Chien, J.-T., 2013. Analysis of driver injury severity in truck-involved accidents using a non-parametric classification tree model. Saf. Sci. 51, 17–22. doi:10.1016/j.ssci.2012.06.017

Chang, L.-Y., Mannering, F., 1999. Analysis of injury severity and vehicle occupancy in truck- and non-truck-involved accidents. Accid. Anal. Prev. 31, 579–592. doi:10.1016/S0001-4575(99)00014-7

Chen, F., Chen, S., 2011. Injury severities of truck drivers in single- and multi-vehicle accidents on rural highways. Accid. Anal. Prev. 43, 1677–1688. doi:10.1016/j.aap.2011.03.026

Chen, P., Shen, Q., 2016. Built environment effects on cyclist injury severity in automobile-involved bicycle crashes. Accid. Anal. Prev. 86, 239–246. doi:10.1016/j.aap.2015.11.002

Cole, G.S., Sherman, A.M., 1995. Microstructural Characterization of Lightweight Structural Materials TransportationLight weight materials for automotive applications. Mater. Charact. 35, 3–9. doi:10.1016/1044-5803(95)00063-1

Daniel, J., Dixon, K., Jared, D., 2000. Analysis of Fatal Crashes in Georgia Work Zones. Transp. Res. Rec. J. Transp. Res. Board 1715, 18–23. doi:10.3141/1715-03

Dong, C., Richards, S.H., Huang, B., Jiang, X., 2015. Identifying the factors contributing to the severity of truck-involved crashes. Int. J. Inj. Contr. Saf. Promot. 22, 116–126. doi:10.1080/17457300.2013.844713

Duncan, C., Khattak, A., Council, F., 1998. Applying the Ordered Probit Model to Injury Severity in Truck-Passenger Car Rear-End Collisions. Transp. Res. Rec. J. Transp. Res. Board 1635, 63–71. doi:10.3141/1635-09

Elghamrawy, T., El-Rayes, K., Liu, L., n.d. Analysis of Injury and Fatal Crashes in Highway Construction Zones, in: Construction Research Congress 2010. American Society of Civil Engineers, pp. 568–577.

Eluru, N., 2013. Evaluating alternate discrete choice frameworks for modeling ordinal discrete variables. Accid. Anal. Prev. 55, 1–11. doi:10.1016/j.aap.2013.02.012

Eluru, N., Bhat, C.R., 2007. A joint econometric analysis of seat belt use and crash-related injury severity. Accid. Anal. Prev. 39, 1037–1049. doi:10.1016/j.aap.2007.02.001

Eluru, N., Bhat, C.R., Hensher, D.A., 2008. A mixed generalized ordered response model for examining pedestrian and bicyclist injury severity level in traffic crashes. Accid. Anal. Prev. 40, 1033–1054. doi:10.1016/j.aap.2007.11.010

FHWA, 2016. Work Zone Facts and Statistics [WWW Document]. URL http://www.ops.fhwa.dot.gov/wz/resources/facts_stats/safety.htm (accessed 6.21.16).

FHWA, 2014. Large Truck and Bus Crash Facts 2012 [WWW Document]. Fed. Mot. Carr. Saf. Adm. URL http://www.fmcsa.dot.gov/safety/data-and-statistics/large-truck-and-bus-crash-facts-2012 (accessed 7.21.15).

FHWA, 2012. Highway Statistics 2012, Highway Statistics Series. FHWA - Office of Highway Policy Information.

FHWA, 2010. Work Zone Facts and Statistics [WWW Document]. URL http://www.ops.fhwa.dot.gov/wz/resources/facts_stats/injuries_fatalities.htm (accessed 7.21.15).

Garber, N., Zhao, M., 2002. Distribution and Characteristics of Crashes at Different Work Zone Locations in Virginia. Transp. Res. Rec. J. Transp. Res. Board 1794, 19–25. doi:10.3141/1794-03

Harb, R., Radwan, E., Yan, X., Pande, A., Abdel-Aty, M., 2008. Freeway Work-Zone Crash Analysis and Risk Identification Using Multiple and Conditional Logistic Regression. J. Transp. Eng. 134, 203–214. doi:10.1061/(ASCE)0733-947X(2008)134:5(203)

Hargroves, B.T., Martin, M.R., 1980. VEHICLE ACCIDENTS IN HIGHWAY WORK ZONES.

Holdridge, J.M., Shankar, V.N., Ulfarsson, G.F., 2005. The crash severity impacts of fixed roadside objects. J. Safety Res. 36, 139–147. doi:10.1016/j.jsr.2004.12.005

Høye, A., 2010. Are airbags a dangerous safety measure? A meta-analysis of the effects of frontal airbags on driver fatalities. Accid. Anal. Prev. 42, 2030–2040. doi:10.1016/j.aap.2010.06.014

Islam, Hernandez, 2013. Large Truck–Involved Crashes: Exploratory Injury Severity Analysis. J. Transp. Eng. 139, 596–604. doi:10.1061/(ASCE)TE.1943-5436.0000539

Islam, M., Hernandez, S., 2013. Modeling Injury Outcomes of Crashes Involving Heavy Vehicles on Texas Highways. Transp. Res. Rec. J. Transp. Res. Board 2388, 28–36. doi:10.3141/2388-05

Khattak, A.J., Khattak, A.J., Council, F.M., 2002. Effects of work zone presence on injury and non-injury crashes. Accid. Anal. Prev. 34, 19–29. doi:10.1016/S0001-4575(00)00099-3

Khattak, A., Targa, F., 2004. Injury Severity and Total Harm in Truck-Involved Work Zone Crashes. Transp. Res. Rec. J. Transp. Res. Board 1877, 106–116. doi:10.3141/1877-12

Khorashadi, A., Niemeier, D., Shankar, V., Mannering, F., 2005. Differences in rural and urban driver-injury severities in accidents involving large-trucks: An exploratory analysis. Accid. Anal. Prev. 37, 910–921. doi:10.1016/j.aap.2005.04.009

Kim, D.-G., Washington, S., 2006. The significance of endogeneity problems in crash models: An examination of left-turn lanes in intersection crash models. Accid. Anal. Prev. 38, 1094–1100. doi:10.1016/j.aap.2006.04.017

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

Li, Y., Bai, Y., 2009. Highway Work Zone Risk Factors and Their Impact on Crash Severity. J. Transp. Eng. 135, 694–701. doi:10.1061/(ASCE)TE.1943-5436.0000055

Li, Y., Bai, Y., 2008a. Development of crash-severity-index models for the measurement of work zone risk levels. Accid. Anal. Prev. 40, 1724–1731. doi:10.1016/j.aap.2008.06.012

Li, Y., Bai, Y., 2008b. Fatal and Injury Crash Characteristics in Highway Work Zones. Presented at the Transportation Research Board 87th Annual Meeting.

Mannering, F.L., Shankar, V., Bhat, C.R., 2016. Unobserved heterogeneity and the statistical analysis of highway accident data. Anal. Methods Accid. Res. 11, 1–16. doi:10.1016/j.amar.2016.04.001

Manski, C.F., McFadden, D. (Eds.), 1981. Structural Analysis of Discrete Data with Econometric Applications. The MIT Press, Cambridge, Mass.

Manual on Uniform Traffic Control Devices (MUTCD), 2009.

McKelvey, R.D., Zavoina, W., 1975. A statistical model for the analysis of ordinal level dependent variables. J. Math. Sociol. 4, 103–120. doi:10.1080/0022250X.1975.9989847

Meng, Q., Weng, J., Qu, X., 2010. A probabilistic quantitative risk assessment model for the long-term work zone crashes. Accid. Anal. Prev. 42, 1866–1877. doi:10.1016/j.aap.2010.05.007

Miller, W.S., Zhuang, L., Bottema, J., Wittebrood, A.J., De Smet, P., Haszler, A., Vieregge, A., 2000. Recent development in aluminium alloys for the automotive industry. Mater. Sci. Eng. A 280, 37–49. doi:10.1016/S0921-5093(99)00653-X

Mooradian, J., Ivan, J.N., Ravishanker, N., Hu, S., 2013. Analysis of driver and passenger crash injury severity using partial proportional odds models. Accid. Anal. Prev. 58, 53–58. doi:10.1016/j.aap.2013.04.022

Nemeth, Z.A., Migletz, D.J., 1978. ACCIDENT CHARACTERISTICS BEFORE, DURING, AND AFTER SAFETY UPGRADING PROJECTS ON OHIO'S RURAL INTERSTATE SYSTEM. Transp. Res. Rec.

Nemeth, Z.A., Rathi, A., 1983. FREEWAY WORK ZONE ACCIDENT CHARACTERISTICS. Transp. Q. 37.

Osman, M., Paleti, R., Mishra, S., Golias, M.M., 2016. Analysis of injury severity of large truck crashes in work zones. Accid. Anal. Prev. 97, 261–273. doi:10.1016/j.aap.2016.10.020

Pahukula, J., Hernandez, S., Unnikrishnan, A., 2015. A time of day analysis of crashes involving large trucks in urban areas. Accid. Anal. Prev. 75, 155–163. doi:10.1016/j.aap.2014.11.021

Pigman, J.G., Agent, K.R., 1990. HIGHWAY ACCIDENTS IN CONSTRUCTION AND MAINTENANCE WORK ZONES. Transp. Res. Rec.

Qi, Y., Srinivasan, R., Teng, H., Baker, R., 2013. Analysis of the frequency and severity of rear-end crashes in work zones. Traffic Inj. Prev. 14, 61–72. doi:10.1080/15389588.2012.675109

Romo, A., Hernandez, S., Cheu, R.L., 2014. Identifying Precrash Factors for Cars and Trucks on Interstate Highways: Mixed Logit Model Approach. J. Transp. Eng. 140, 04013016. doi:10.1061/(ASCE)TE.1943-5436.0000621

Rouphail, N.M., Yang, Z.S., Fazio, J., 1988. COMPARATIVE STUDY OF SHORT- AND LONG-TERM URBAN FREEWAY WORK ZONES. Transp. Res. Rec.

Saltzman, G.M., Belzer, M.H., 2007. Truck Driver Occupational Safety and Health (Conference Report No. DHHS (NIOSH) Publication No. 2007–120). National Institute for Occupational Safety and Health (NIOSH).

Savolainen, P., Mannering, F., 2007. Probabilistic models of motorcyclists' injury severities in single- and multi-vehicle crashes. Accid. Anal. Prev. 39, 955–963. doi:10.1016/j.aap.2006.12.016

Schrock, S.D., Ullman, G.L., Cothron, A.S., Kraus, E., Voigt, A.P., 2004. AN ANALYSIS OF FATAL WORK ZONE CRASHES IN TEXAS.

Schwarz, G., 1978. Estimating the Dimension of a Model. Ann. Stat. 6, 461–464. doi:10.1214/aos/1176344136

Shankar, V., Mannering, F., 1996. An exploratory multinomial logit analysis of single-vehicle motorcycle accident severity. J. Safety Res. 27, 183–194. doi:10.1016/0022-4375(96)00010-2

Shankar, V., Mannering, F., Barfield, W., 1996. Statistical analysis of accident severity on rural freeways. Accid. Anal. Prev. 28, 391–401. doi:10.1016/0001-4575(96)00009-7

Ulfarsson, G.F., Mannering, F.L., 2004. Differences in male and female injury severities in sport-utility vehicle, minivan, pickup and passenger car accidents. Accid. Anal. Prev. 36, 135–147. doi:10.1016/S0001-4575(02)00135-5

Ullman, G.L., Ullman, B.R., Finley, M.D., 2006. Analysis of Crashes at Active Night Work Zones in Texas. Presented at the Transportation Research Board 85th Annual Meeting.

U.S. Department of Transportation, 2014. Traffic Safety Facts 2012 (No. DOT HS 812 032), Traffic Safety Fact Annual Report. National Highway Traffic Safety Administration (NHTSA).

Wang, H.L., Shi, Z.K., 2013. Hierarchical Regression Model for Truck Collision Severity Analysis. Adv. Mater. Res. 671-674, 2889–2892. doi:10.4028/www.scientific.net/AMR.671-674.2889

Wang, Z., Lu, J., Wang, Q., Lu, L., Zhang, Z., 2010. Modeling Injury Severity in Work Zones using Ordered PROBIT Regression, in: ICCTP 2010. American Society of Civil Engineers, pp. 1058–1067.

Weng, J., Meng, Q., 2011. Analysis of driver casualty risk for different work zone types. Accid. Anal. Prev. 43, 1811–1817. doi:10.1016/j.aap.2011.04.016

Wong, J.M., Arico, M.C., Ravani, B., 2011. Factors Influencing Injury Severity to Highway Workers in Work Zone Intrusion Accidents. Traffic Inj. Prev. 12, 31–38. doi:10.1080/15389588.2010.525569

Yasmin, S., Eluru, N., 2013. Evaluating alternate discrete outcome frameworks for modeling crash injury severity. Accid. Anal. Prev. 59, 506–521. doi:10.1016/j.aap.2013.06.040

Yasmin, S., Eluru, N., Ukkusuri, S.V., 2014. Alternative Ordered Response Frameworks for Examining Pedestrian Injury Severity in New York City. J. Transp. Saf. Secur. 6, 275–300. doi:10.1080/19439962.2013.839590

Ye, F., Lord, D., 2014. Comparing three commonly used crash severity models on sample size requirements: Multinomial logit, ordered probit and mixed logit models. Anal. Methods Accid. Res. 1, 72–85. doi:10.1016/j.amar.2013.03.001

Zhu, X., Srinivasan, S., 2011a. A comprehensive analysis of factors influencing the injury severity of large-truck crashes. Accid. Anal. Prev. 43, 49–57. doi:10.1016/j.aap.2010.07.007

Zhu, X., Srinivasan, S., 2011b. Modeling occupant-level injury severity: An application to large-truck crashes. Accid. Anal. Prev. 43, 1427–1437. doi:10.1016/j.aap.2011.02.021