AN ABSTRACT OF THE DISSERTATION OF

Nabeel Saleem Saad Al-Bdairi for the degree of Doctor of Philosophy in Civil Engineering presented on June 29, 2018.

Title: Modeling Unobserved Heterogeneity and the Injury Severities of Truck Drivers in Run-Off-Road (ROR) Crashes: Econometric Methods and Applications

Abstract approved:

______________________________________________________
Salvador Hernandez

Recent statistics regarding large truck crashes reveal that fatality rates of large trucks per 100 million vehicle miles traveled (VMT) and fatality rates per 1,000 registered vehicles are higher than those for passenger vehicles. These statistics underscore the need for greater efforts by safety professionals to help mitigate the impacts of these types of crashes on society and on the ground freight transportation industry as a whole. Of special interest are run-off-road (ROR) crashes (also referred to as roadway departure), which are crashes that occur due to a vehicle crossing an edge line or a center line of a roadway or/and leaving the designated lane. These types of crashes roughly constituted 54% of all traffic fatalities in the U.S. for the period between 2013 and 2015. There have been several efforts that have addressed large truck-involved crashes from varying perspectives (e.g., time of day, facility type, and single vs. multi-vehicle, work zone). However, there are several research gaps that need further attention, for example, better understanding of the relationship between contributing factors and driver injury severities due to ROR crashes involving large trucks. Specifically, those driver injury severities related to the impact of lighting conditions (dark vs. lighted) and land use (urban vs. rural). From a methodological modeling perspective accounting for unobserved heterogeneity has traditionally been ignored in regard to ROR driver injury severity analyses. To address these gaps, this dissertation aims to develop and estimate advanced econometric models that account for
unobserved heterogeneity to better understand the contributing factors of driver injury severity of ROR crashes involving large trucks in the state of Oregon.

This dissertation includes three manuscripts that investigate injury severity of large truck drivers involved in ROR crashes in the state of Oregon for the period 2007 to 2014. In the first manuscript, an ordered random parameter probit model was estimated to predict the likelihood of three injury severity categories using Oregon crash data: severe injury (fatal and incapacitating), minor injury (non-incapacitating and possible injury), and no injury while addressing the unobserved heterogeneity. The modeling framework presented in this manuscript offers a flexible methodology to analyze ROR crashes involving large trucks while accounting for unobserved heterogeneity.

The second manuscript examines the impact of lighting conditions on injury severity of large truck drivers involved in ROR crashes. This was done by disaggregating crash data by lighting conditions into two datasets: one for the lighted condition and the other for the dark condition. Hence, two separate mixed logit models were developed to capture the contributing factors that affect injury severity in each lighting condition while accounting for unobserved heterogeneity. To validate the estimation results, series of likelihood ratio tests were conducted. Model separation tests along with estimation results indicate that lighting conditions need to be analyzed separately with 99.99% confidence.

Lastly, an in-depth analysis was conducted in the third manuscript to examine the effect of land use setting (urban vs. rural) on injury severity of large truck drivers involved in ROR crashes. Again, disaggregating crash data was achieved to create two independent datasets: one pertaining to ROR crashes involving large trucks occurring on urban and the other for those occurring on rural areas. Instead of utilizing random parameter approach as a framework to account for unobserved heterogeneity, this manuscript utilizes two latent class ordered probit models to capture factors exclusively that contribute to each land use type. Once again, model separation tests along with estimation results reveal that there are distinctions in terms of contributing factors based on land use type. Therefore, ROR crashes involving large trucks need to be analyzed separately based on land use with 99.99% confidence.
It is expected that estimating advanced econometric methods to identify contributing factors to injury severity of large truck drivers involved in ROR crashes while accounting for unobserved heterogeneity can be used as a basis to aid transportation safety engineers, trucking industry, transportation planners, and state agencies in implementing appropriate safety countermeasures to help mitigate ROR crashes. In terms of study implications, the estimation results of this dissertation have direct implications on safety policy. For instance, the finding of this dissertation reinforces the notion that the current policy regarding seatbelt usage should be amended in an attempt to increase seatbelt usage by penalizing drivers violating this policy to deterrent fines and forcing them to enroll in a required defensive driving course in the state of Oregon.
Modeling Unobserved Heterogeneity and the Injury Severities of Truck Drivers in Run-Off-Road (ROR) Crashes: Econometric Methods and Applications

by
Nabeel Saleem Saad Al-Bdairi

A DISSERTATION

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Doctor of Philosophy

Presented June 29, 2018
Commencement June 2019
Doctor of Philosophy dissertation of Nabeel Saleem Saad Al-Bdairi presented on June 29, 2018

APPROVED:

Major Professor, representing Civil Engineering

Head of the School of Civil and Construction Engineering

Dean of the Graduate School

I understand that my dissertation will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my dissertation to any reader upon request.

Nabeel Saleem Saad Al-Bdairi, Author
ACKNOWLEDGEMENTS

Foremost, I would like to sincerely thank my advisor Dr. Salvador Hernandez for his unconditional support and invaluable insight in this dissertation in my Ph.D. pursuit. This dissertation would not have been possible without the invaluable advice, interest, and contribution of him. His enthusiasm, encouragement, and faith in me is extremely inspiring and helped shaped this dissertation. He was always available for my non-stopping questions and he was positive and gave a great deal of his time and vast expertise. Undoubtedly, he is not an academic advisor that provided insightful advice, but he also served as an excellent life mentor to me. My academic attitude as a researcher and a university professor has been positively and significantly influenced by the way that Dr. Hernandez follows with his students, which is summarized by bringing a balanced dosage of work and fun to the table so that everything flows so effortlessly.

I also would like to extend my special thanks to the members of my dissertation committee: Dr. Chris A. Bell, Dr. David S. Hurwitz, Dr. Erdem Coleri, and Dr. Hector Vergara for their encouragement and insightful comments. Definitely, these valuable suggestions have immensely contributed to further improve the quality of this dissertation.

Moreover, I would like to thank The Higher Committee for Education Development (HCED) in Iraq for providing a scholarship to pursue my Ph.D. at Oregon State University. Finally, I would like to thank my family: sisters, brothers, wife, and children (Hayder, Zain, and Adam) for their absolute support throughout my Ph.D. journey. Above all, I would like to thank my wife for her patience, encouragement, unending sacrifice, and support. None of these would have been possible without her. Thanks for being present in my life, all the time, all the way!!

Last but not least, my acknowledgment would be incomplete without thanking my parents. This dissertation is dedicated to my Mom and my late father (Saleem Al-Bdairi) who painstakingly sacrificed all they had to shape my future to help me excel in my life. I would not have been where I am today and what I am today without your
love and unconditional support. I would like to confess that you are the only ones who can love and give without expecting anything in return. Being present in this world surrounded by your parent is a blessing that requires gratitude to God for this grace, which you will not feel unless you lose it. Thank you, dad! I know you are up there, listening, watching over me, and sending me your blessings constantly.

This achievement is for you dad!
CONTRIBUTION OF AUTHORS

Dr. Salvador Hernandez provided extensive feedback on three manuscripts and his assistance with the data processing and the daily basis insightful recommendations was a huge part of making this dissertation a successful project. Jason Anderson helped edit the second manuscript for English grammar and structure. Also, he helped write a part of summary and conclusions in the second manuscript section.
TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter 1: Introduction</th>
<th>.................................................................</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1 Motivation</td>
<td>..................................................................................</td>
<td>1</td>
</tr>
<tr>
<td>1.2 Objectives</td>
<td>..................................................................................</td>
<td>4</td>
</tr>
<tr>
<td>1.3 Contribution of Dissertation</td>
<td>..................................................................................</td>
<td>6</td>
</tr>
<tr>
<td>1.4 Organization of Dissertation</td>
<td>..................................................................................</td>
<td>7</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chapter 2: An empirical analysis of run-off-road injury severity crashes involving large trucks</th>
<th>..................................................................................</th>
<th>19</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1 Introduction</td>
<td>..................................................................................</td>
<td>21</td>
</tr>
<tr>
<td>2.2 Data Description</td>
<td>..................................................................................</td>
<td>24</td>
</tr>
<tr>
<td>2.3 Methodological Approach</td>
<td>..................................................................................</td>
<td>27</td>
</tr>
<tr>
<td>2.3.1 Ordered Random Parameter Probit Model</td>
<td>..................................................................................</td>
<td>28</td>
</tr>
<tr>
<td>2.4 Results and Discussion</td>
<td>..................................................................................</td>
<td>31</td>
</tr>
<tr>
<td>2.4.1 Environmental and Roadway-Related Factors</td>
<td>..................................................................................</td>
<td>35</td>
</tr>
<tr>
<td>2.4.2 Human-Related Factors</td>
<td>..................................................................................</td>
<td>36</td>
</tr>
<tr>
<td>2.4.3 Vehicle and Crash-Related Factors</td>
<td>..................................................................................</td>
<td>38</td>
</tr>
<tr>
<td>2.5 Conclusions and Future Research</td>
<td>..................................................................................</td>
<td>39</td>
</tr>
</tbody>
</table>
TABLE OF CONTENTS (Continued)

<table>
<thead>
<tr>
<th>Chapter 3 : Contributing Factors to Run-off-Road Crashes Involving Large Trucks under Lighted and Dark Conditions</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1 Introduction</td>
<td>48</td>
</tr>
<tr>
<td>3.2 Literature Review</td>
<td>49</td>
</tr>
<tr>
<td>3.3 Methodological Approach</td>
<td>51</td>
</tr>
<tr>
<td>3.4 Empirical Setting</td>
<td>54</td>
</tr>
<tr>
<td>3.4.1 Model Separation</td>
<td>57</td>
</tr>
<tr>
<td>3.5 Estimation Results</td>
<td>60</td>
</tr>
<tr>
<td>3.5.1 Dark Conditions</td>
<td>63</td>
</tr>
<tr>
<td>3.5.2 Lighted Conditions</td>
<td>67</td>
</tr>
<tr>
<td>3.6 Summary and Conclusions</td>
<td>70</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chapter 4 : Urban and rural run-off-road (ROR) crashes and the injury severities of large truck drivers</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1 Introduction</td>
<td>78</td>
</tr>
<tr>
<td>4.2 Literature Review</td>
<td>80</td>
</tr>
<tr>
<td>4.3 Data Description</td>
<td>83</td>
</tr>
<tr>
<td>4.4 Methodology</td>
<td>87</td>
</tr>
</tbody>
</table>
TABLE OF CONTENTS (Continued)

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.4.1 Model separation tests</td>
<td>90</td>
</tr>
<tr>
<td>4.5 Estimation results</td>
<td>92</td>
</tr>
<tr>
<td>4.5.1 Factors contributing to ROR crashes on both models</td>
<td>97</td>
</tr>
<tr>
<td>4.5.2 Factors contributing to ROR crashes on urban roadways</td>
<td>102</td>
</tr>
<tr>
<td>4.5.3 Factors contributing to ROR crashes on rural roadways</td>
<td>104</td>
</tr>
<tr>
<td>4.6 Summary and Conclusions</td>
<td>105</td>
</tr>
<tr>
<td>Chapter 5 : Conclusions and Recommendations</td>
<td>112</td>
</tr>
<tr>
<td>5.1 Summary and Conclusions</td>
<td>112</td>
</tr>
<tr>
<td>5.2 Key findings</td>
<td>114</td>
</tr>
<tr>
<td>5.3 Research Implications</td>
<td>116</td>
</tr>
<tr>
<td>5.4 Practical Applications</td>
<td>118</td>
</tr>
<tr>
<td>5.5 Limitations and Directions for Future Research</td>
<td>118</td>
</tr>
<tr>
<td>Chapter 6 : References</td>
<td>121</td>
</tr>
</tbody>
</table>
LIST OF TABLES

Table 1.1: Summary of previous studies analyzing injury severities of ROR crashes . 9
Table 2.1: Frequency and percentage distribution of driver injury .................................. 26
Table 2.2: Estimated Results for the Ordered Probit Model with Fixed and Random Parameters (Values in Parentheses Indicate the Standard Deviation of the Random Parameters Distribution) .................................................................................. 32
Table 2.3: Marginal Effects for the Ordered Probit Model with Random Parameters 33
Table 3.1: Frequency Distribution of the Selected Variables under Dark Conditions 56
Table 3.2: Frequency Distribution of the Selected Variables under Lighted Conditions .................................................................................................................. 57
Table 3.3: Mixed Logit Estimation Results for Dark Conditions (Values in Parentheses Indicate the Standard Deviation of the Random Parameters Distribution) ........................................................................................................ 61
Table 3.4: Mixed Logit Estimation Results for Lighted Conditions (Values in Parentheses Indicate the Standard Deviation of the Random Parameters Distribution) .............................................................................................. 62
Table 4.1: Frequency and percentage distribution of driver injury in urban and rural 84
Table 4.2: Description of selected explanatory variables used in the analyses of urban and rural models ............................................................................................................. 85
Table 4.3: Frequency and percentage distribution of explanatory variables in urban model .................................................................................................................................. 86
Table 4.4: Frequency and percentage distribution of explanatory variables in rural model .................................................................................................................................. 86
Table 4.5: Estimated probabilities for each latent class in urban and rural models .... 93
Table 4.6: Estimation results of latent class ordered probit for urban and rural models .................................................................................................................................. 95
Table 4.7: Estimated marginal effects for latent class ordered probit of urban model96
## LIST OF TABLES (Continued)

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 4.8: Estimated marginal effects for latent class ordered probit of rural model</td>
<td>96</td>
</tr>
</tbody>
</table>
Chapter 1: Introduction

1.1 Motivation

Traffic safety is a major concern for road users and governmental agencies that plan, construct, and maintain the highway system. As such, there is a growing concern and interest in studying traffic safety. This attention toward traffic safety is stemming from the devastating impacts of roadway crashes on the society. Roadway crashes are a serious and a global health problem that needs to be thoroughly addressed because more than 1.2 million people die in the world annually due to roadway crashes, indicating that roadway crashes are the ninth leading cause of death across all age groups (WHO, 2015). Unfortunately, those crashes are expected to be the seventh leading cause of death for individuals of all ages by 2030 (WHO, 2015). Those crashes are, generally, accompanied by catastrophic economic and emotional burden on society. As regards to economic losses, Blincoe et al. (2015) estimated the economic burden of roadway crashes occurred in the U.S in 2010 by $242 billion (equivalent to 1.6% of the U.S. GDP). This estimate included costs of a wide range of losses such as medical treatments, legal and court costs, emergency service, insurance administration, congestion costs, property damage, workplace losses, and losing productivity. Turning to societal effects, Chang and Mannering (1999) demonstrated that the influences of crashes on society and involved people could be represented by the pain and suffering that individuals would incur, medical costs, wage loss, higher insurance premium rates, and vehicle repair costs. Given these statistics, there are substantial needs to better understand significant contributing factors and their impacts on injury severities of roadway crashes.

Large truck-involved crashes are of particular interest to the transportation officials, the trucking industry, and transportation researchers. FARS defines a large truck as a truck with a gross vehicle weight rating (GVWR) greater than 10,000 pounds (FMCSA, 2016). The trucking industry has significant contributions to the U.S. economy by moving and hauling goods throughout the country. In 2013, large trucks moved a daily average of about 55 million tons of freight valued at more than $49.3 billion in the U.S.
Unfortunately, the crucial role that large trucks play in the U.S. economy does not come without a price. This price is represented by roadway crashes and resulting fatalities that result due to the increase in the number of registered large trucks. According to recent statistics released by FMCSA (2016), approximately 35,092 people killed in crashes involving all vehicles in the U.S. in 2015. In the same year, total fatalities in large truck crashes accounted for 11.6% (4067 fatalities) (FMCSA, 2016). These statistics regarding large trucks reveal that one out of 10 fatalities occurred due to large truck-involved crashes.

The interest underlying large truck safety stems from three reasons. First, the cost of large truck-involved crashes is high compared to crashes involving passenger cars in the sense that these crashes pose a significant impact on the economy. Zaloshnja and Miller (2007) based on 2005 dollars, provided estimates of large trucks involved crashes by crash severity. In their study, fatal crashes, non-fatal crashes, and PDO crashes were estimated as: $3,604,518, $195,258, and $15,114, respectively. These costs encompass medical costs, emergency services costs, property damage costs, lost productivity, and the monetized value of the pain, suffering, and quality of life lost because of a death or injury (Zaloshnja and Miller, 2007). Second, FMCSA (2016) reveals in their recent statistics regarding large truck crashes that large trucks due to its operating characteristics such as weight, size, length, high mileage exposure, and poor acceleration/deceleration performance pose much higher risk compared to passenger vehicles by many measures. For example, even though large trucks accounted for 4% of all registered vehicles and 9% of the total vehicle miles traveled (VMT) in 2016, the fatality rates of large trucks per 100 million vehicle miles traveled (VMT) are higher than those for passenger vehicles for 10 years period between 2006 and 2015 (NHTSA, 2018). In terms of another comparative measure, fatality rates per 1000 registered vehicles, each individual large truck was at least two times more likely to be involved in a fatal crash as opposed to a passenger vehicle for the same time period (FMCSA, 2016). Third, at the state level, the state of Oregon experienced a significant increasing trend in total crashes, truck-related crashes, and fatalities associated with such crashes, particularly after the recovering of the national economy from the recession in 2008.
These alarming trends in truck-related crashes and its fatalities underscore the importance of further studies in large truck safety. Therefore, it is anticipated that investment in large truck safety (in the state of Oregon in particular) would be greatly beneficial to society from a return-on-investment perspective in the sense that transportation agencies should focus their efforts to invest more in the safety of trucks (Knipling, 2015).

Over the last few years considerable attention has been placed on addressing the issues related to injury severity sustained by the drivers of large trucks. Most of this attention has revolved around understanding the contributing factors to the various degrees of injury severity sustained by large truck drivers from various perspectives (e.g., time of day, facility type, and single vs. multi-vehicle, work zone) (Anderson and Hernandez, 2017; Chang and Chien, 2013; Chen and Chen, 2011; Islam and Hernandez, 2013a, 2013b; Islam et al., 2014; Khattak and Targa, 2004; Khorashadi et al., 2005; Lemp et al., 2011; Naik et al., 2016; Pahukula et al., 2015; Zhu and Srinivasan, 2011). What is still not clearly understood are the relationship between the injury severity sustained by drivers of large trucks, lighting conditions (dark vs. lighted)), and land use type (urban vs. rural) on contributing factors for Run-off-road crashes (ROR). Run-off-road (ROR) (also known as roadway departure) crashes according to FMCSA (2017), are crashes that occur due to a vehicle crossing an edge line or a center line of a roadway or/and leaving the designated lane. These types of crashes roughly constituted 54% of all traffic fatalities in the U.S. for the period between 2013 and 2015. A similar trend has been observed in the state of Oregon in which nearly 55% of all fatalities between 2009 and 2015 were due to ROR crashes (Federal Highway Administration, 2017).

In order to address the problem of ROR crashes involving large trucks in the sense of saving people lives and reducing injuries resulting from these crashes, identifying contributing factors to ROR crashes is paramount. Yet, this is not an easy task because the crash data used for analyses are usually obtained from police crash reports, and these reports are characterized by some drawbacks. For example, police crash reports suffer from underreporting bias in the sense that drivers tend to not report crashes with low severity or PDO crashes, especially crashes in which they were intoxicated and/or
guilty. As such, less severe and/or PDO crashes are more likely to be underreported (Mannering and Bhat, 2014; Yamamoto et al., 2008; Ye and Lord, 2011). Another issue that arises in police crash reports is the discrepancies in injury severity recorded in these reports compared to hospital reports for the same crashes (Mannering and Bhat, 2014). This variation in injury severity reporting between two data sources can be related to the failure of drivers to provide accurate details due to their mental state at that time of incident (Shinar, 2007). These shortcomings in police reporting, in addition to a few others, for example, an officer’s state of mind while collecting and reporting a crash introduces unobserved heterogeneity (also known as unobserved factors). Here, unobserved heterogeneity refers to those factors and information that are not included in the crash reports and/or are latent to the analysts. In terms of the effect of unobserved factors on injury severity of traffic crashes, these factors have a variation in their influences rather than having fixed effects (Mannering et al., 2016; Washington et al., 2011). All these drawbacks in crash data if not appropriately remediated, would lead to erroneous inferences and/or under- or over-estimation in the injury severity sustained by drivers.

With this in mind, there is a need to address the injury severities of drivers of large trucks for ROR crashes since these crashes comprise a large portion of crashes that lead to higher injury severity sustained by truck drivers. Furthermore, there is a growing interest to better understand the impacts that lighting conditions and land use setting have on either increasing or decreasing the injury severities sustained while taking into account unobserved heterogeneity.

1.2 Objectives

Although there have been several efforts that address large truck-involved crashes from varying perspectives (e.g., time of day, facility type, and single vs. multi-vehicle, work zone as mentioned earlier), the relationship between contributing factors and severity of ROR crashes involving large trucks needs to be addressed. Injury severity of ROR crashes has been examined in previous efforts as illustrated in Table 1.1. Still, ROR crashes in these studies have historically been introduced in modeling
frameworks as an indicator variable. These indicators variables are valuable in gaining an overall appreciation of how injury severities may vary given their inclusion in the modeling framework. Yet, this approach is limited and does not allow for a clearer understanding of the specific factors that play a role in injury crashes under ROR conditions. In addition to studying the effect of ROR crashes on injury severities, lighting conditions (dark vs. lighted) and land use setting (urban vs. rural) distinctions will be examined to study their effect on ROR injury severities. These distinctions have been studied in the literature, but have not been studied in the context of ROR crashes and their impact on injury severities of large truck-involved crashes. Furthermore, past studies have primarily focused on ROR crashes without any vehicle type distinction (i.e., considered all vehicle types). From a methodological perspective, past studies have not accounted for the presence of unobserved heterogeneity.

In light of these drawbacks in literature, this study focuses on better understanding the relationship between contributing factors and the injury severity of ROR crashes involving large trucks, which are still not clearly understood. To do so, unobserved heterogeneity methods will be examined and tested (random parameters ordered probit model, mixed logit model, and latent class ordered probit model). That is, the current study considers the variation in the nature and extent of injury severity of ROR crashes depending on lighting condition and land use settings. As such, these crashes are modeled in this study separately based on subpopulations (land use and lighting condition) rather than holistically to determine their impact on driver injury severities. This is accomplished by utilizing advanced econometric modeling frameworks to account for the limitations present in the police crash reports through unobserved heterogeneity extensions of these econometric models. Hence, the current study seeks to fill the gap in the literature through analyzing driver injury severity of ROR crashes involving large trucks under lighting conditions (dark vs. lighted) and land use (urban vs. rural) distinctions. This is accomplished through the following aims:

- How do environmental conditions, roadway characteristics, driver characteristics, and vehicle factors affect ROR crash injury severity outcomes?
• How do lighting conditions (dark vs. lighted) influence injury severities of ROR large truck-involved crashes? What are the exclusive contributing factors for each lighting condition, and what are the commonalities? Is this distinction valid?

• How does land use (urban vs. rural) influence injury severities of ROR large truck-involved crashes? What are the exclusive contributing factors for each land use setting, and what are the commonalities? Is this distinction valid?

Lastly, the findings of this study can provide valuable information that can aid state agencies, safety planners, transportation safety researchers, and the trucking industry in identifying appropriate countermeasures specific to ROR crashes involving the drivers of large trucks that occur in urban, rural, lighted, and dark conditions. By doing so, the number and severity of ROR crashes involving large trucks can be reduced, in turn, saving lives and societal costs. To the best of the authors’ knowledge, these are the first attempts to comprehensively analyze injury severity of ROR crashes involving the drivers of large trucks through disaggregated levels by disaggregating crash data based on land use setting (rural vs. urban) and lighting condition (lighted vs. dark) by developing advanced econometric models.

1.3 Contribution of Dissertation

This dissertation contributes to the current state-of-the-art both empirically and methodologically. From an empirical perspective, the current dissertation has primarily focused on better understanding the contributing factor for ROR crashes involving large trucks. Studies involving ROR crashes are not new, however studies that research ROR involving large trucks are scarce and have primarily focused on impacts of the event on injury severities of drivers rather than identifying the factors for the occurrences of ROR crashes. This dissertation uniquely captures this aspect by addressing the contributing factors to ROR crashes. These factors include but are not limited to driver characteristics, roadway characteristics, vehicle factors,
environmental factors, and crash characteristics. In addition to identifying contributing factors to ROR involving large trucks, this dissertation estimates marginal effects that quantify the impacts of each explanatory variable on the injury severity sustained by the drivers. In doing so, the analyst can better identify potential countermeasures based on influence of the factors on ROR crashes. This dissertation further extends the study of ROR crashes involving large trucks by investigating the effects of lighting conditions and land use setting on injury severity sustained by the drivers of large trucks. These extensions have not been applied to ROR crashes and is a first in the context of studying the injury severities sustained of drivers of large truck. In considering these extensions, this dissertation contributes to the context of large truck safety by comprehensively analyzing the effect of lighting conditions and land use setting at disaggregate levels.

Turning to methodological contributions, previous studies have extensively studied ROR crashes (with the main focus on passenger vehicles or vehicular mix). Past studies are characterized by some limitations. For instance, a key shortcoming of past studies is that unobserved heterogeneity (unobserved factors in the crash data) has not been addressed. As such, drawn inferences from past studies can be misleading and erroneous inferences drawn. Such limitation has been mainly accounted for in this dissertation through the utilization of two methods to account for unobserved heterogeneity, namely the random parameter and latent class approaches. Accounting for unobserved heterogeneity allows an analyst to better infer on the impact of specific contributing factors and provides a mechanism for capturing the impacts of unobservable factors on ROR crashes involving large trucks.

1.4 Organization of Dissertation

This dissertation is a compilation of three journal manuscripts that aims to holistically examine injury severity sustained by drivers of large trucks involved in ROR crashes. Chapter 2 represents the first manuscript that has been published in “Accident Analysis and Prevention” in which a random parameters ordered probit model was estimated to capture the contributing factors to the injury severity sustained
by drivers involved in ROR large truck crashes and to account for unobserved heterogeneity present in the crash dataset (Al-Bdairi and Hernandez, 2017). Chapter 3 represents the second manuscript that has been published in “Journal of Transportation Engineering, Part A: Systems”. This chapter discusses investigating injury severity of ROR crashes in depth by splitting crash data pertaining to large trucks in the state of Oregon by lighting conditions into two datasets: one for lighted conditions and the other for the dark conditions. Injury severities of ROR crashes involving the drivers of large trucks were modeled separately by lighting conditions rather than holistically to further capture factors exclusively pertaining to injury severity sustained in each lighting condition and to account for unobserved heterogeneity. To validate the estimation results, a series of likelihood ratio tests were conducted (Al-Bdairi et al., 2018). Chapter 4 illustrates the third manuscript that discusses an alternative approach to unobserved heterogeneity present in the crash data, namely the latent class ordered probit models. This chapter presents the differences and commonalities in injury severity of ROR crashes involving large trucks occurring in urban and rural settings. The final chapter, chapter 5, discusses major findings and conclusions of the current study, illustrates the practical applications of these findings, and highlights limitations in the current study and suggests some directions for future research.
Table 1.1: Summary of previous studies analyzing injury severities of ROR crashes

<table>
<thead>
<tr>
<th>Study</th>
<th>Data used</th>
<th>Methodological approach</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Gong and Fan, 2017)</td>
<td>5-year crash data pertaining to driver injury severity involved in single-vehicle ROR crashes occurring between 2009 and 2013 in North Carolina were used. Injury severity has been categorized into five injury levels: fatality (F), incapacitating injury (I), non-incapacitating injury (N), possible injury (P) and PDO.</td>
<td>Mixed logit model</td>
<td>Despite the authors addressed the unobserved heterogeneity in this study. Yet, this study has failed to disaggregate crash data based on crash type. Instead, disaggregation has been done by driver age into three age groups: young (ages 16–24), middle-aged (ages 25–65), and older drivers (ages over 65). Further, only crashes occurring on rural roadways were considered. Log-likelihood ratio tests that validate this disaggregation has not been provided.</td>
</tr>
<tr>
<td>(Gong et al., 2016)</td>
<td>5-year crash data pertaining to driver injury severity involved in single-vehicle ROR crashes occurring between 2009 and 2013 in North Carolina were used. Injury severity has been categorized into five injury levels: fatality (F), incapacitating injury (I), non-incapacitating injury (N), possible injury (P) and PDO.</td>
<td>Four different discrete choice models were used: traditional ordered logit model (ORL), multinomial logit model (MNL), partial proportional odds model (PPO), and mixed logit model (ML)</td>
<td>The authors have conducted several discrete choice models, including ordered and unordered data. Also, the main focus of this study is ROR crashes involving a single vehicle. However, the distinction between vehicles was neglected. Further, only crashes occurring on rural roadways were considered. Lastly, the log-likelihood ratio tests that validate this disaggregation has not been provided.</td>
</tr>
<tr>
<td>(Shawky et al., 2016)</td>
<td>The data used were obtained from Abu Dhabi crash reports in UAE for ROR crashes occurred between 2007 and 2013. Injury severity was treated as ordered in nature with four levels: fatal, severe injury, medium injury, and severe injury.</td>
<td>An ordered probit model was applied because injury severity is ordered in nature.</td>
<td>This study failed to account for unobserved heterogeneity in crash data because fixed parameters ordered probit model was used. Moreover, a holistic model was conducted. Further, marginal effects were not used for interpretation of findings.</td>
</tr>
</tbody>
</table>
Table 1.1: Summary of previous studies analyzing injury severities of ROR crashes (Continued)

<table>
<thead>
<tr>
<th>Study</th>
<th>Data used</th>
<th>Methodological approach</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Roque et al., 2015)</td>
<td>ROR crashes that occurred on Portuguese freeways between 2009 and 2010 were obtained from police crash reports. A total of 764 ROR crashes were considered for analyses in which injury severity was grouped into four main groups: fatal injury, severe injury, minor injury, and property damage only.</td>
<td>Discrete choice models, namely multinomial logit model and mixed logit model were used.</td>
<td>This study addressed heterogeneity issue in crash data. However, the sample size of the crash data used was very small that arises the issue of incorrect parameter estimates. Furthermore, only ROR crashes involving passenger car were considered in the analyses.</td>
</tr>
<tr>
<td>(Montella et al., 2015)</td>
<td>ROR crashes occurred between 2001 and 2011 on A16 in Italy were obtained from police reports. Injury severity was classified into three categories (property damage only, slight injury, and severe injury/fatal).</td>
<td>A multinomial logit model was conducted.</td>
<td>Again, heterogeneity in crash data was ignored. Therefore, this issue can lead one to draw erroneous inferences.</td>
</tr>
<tr>
<td>(Amarasingha and Dissanayake, 2014)</td>
<td>This study obtained crash data concerning ROR and non-ROR crashes occurred between 2007 and 2011 in the state of Kansas from the Kansas Department of Transportation (KDOT). Injury severity was used as a binary variable.</td>
<td>Binary logistic regression models were developed.</td>
<td>Grouping injury severity into a binary outcome is considered a simple starting method in injury severity analyses. Moreover, the inferences drawn will be erroneous because heterogeneity was ignored. Further, the study did not consider a particular vehicle type or any disaggregation.</td>
</tr>
<tr>
<td>(Eustace et al., 2014)</td>
<td>The data used were obtained from Ohio Department of Public Safety (ODPS) crash database. These data included all ROR crashes occurred in the state of Ohio between 2008 and 2012. Injury severity was categorized into three levels: no injury, possible injury and non-incapacitating injury, and incapacitating injury and fatal injury.</td>
<td>For exploratory analysis, the decision tree model was used, while generalized ordered logit model was developed to identify factors that increase the probability of involving in ROR related crashes.</td>
<td>The analyses conducted in this study were exploratory in nature by using decision tree model and using generalized ordered logit model for injury severity of ROR crashes. This study treated factors as fixed across observations in the sense that heterogeneity in crash data was ignored. Moreover, the authors did not consider any disaggregation of data by subpopulations.</td>
</tr>
</tbody>
</table>
Table 1.1: Summary of previous studies analyzing injury severities of ROR crashes (Continued)

<table>
<thead>
<tr>
<th>Study</th>
<th>Data used</th>
<th>Methodological approach</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Dissanayake and Roy, 2014)</td>
<td>ROR crashes occurred between 1999 and 2008 in the state of Kansas were obtained from Kansas accident reporting system (KARS).</td>
<td>Binary logistic regression models were used.</td>
<td>Unobserved heterogeneity was ignored in the analyses. Moreover, the authors failed to identify factors exclusively contribute to each type of vehicle involved or land use setting.</td>
</tr>
<tr>
<td>(Eustace et al., 2014)</td>
<td>The data used were obtained from Ohio Department of Public Safety (ODPS) crash database. These data included all ROR crashes occurred in the state of Ohio between 2008 and 2012. Injury severity was categorized into three levels: no injury, possible injury and non-incapacitating injury, and incapacitating injury and fatal injury.</td>
<td>For exploratory analysis, the decision tree model was used, while generalized ordered logit model was developed to identify factors that increase the probability of involving in ROR related crashes.</td>
<td>The analyses conducted in this study were both exploratory by using decision tree model and using generalized ordered logit model for injury severity incurred by ROR crashes. In both cases, the approaches utilized treated factors as fixed across observations in the sense that heterogeneity in crash data was ignored. Moreover, the sample size of crash data used was very large containing more than 384,000 ROR crashes. Yet, the authors did not consider any disaggregation of data by subpopulations.</td>
</tr>
<tr>
<td>(Peng and Boyle, 2012)</td>
<td>The crash data used were obtained from Washington State Department of Transportation (WSDOT) for the period between the years 2006 and 2009 for ROR crashes involving commercial vehicles. The injury severity was treated as a binary outcome with 1 representing (injury and fatal) and 0 for PDO.</td>
<td>In this study, a binary logistic regression model was used.</td>
<td>Grouping injury severity into a binary outcome is considered a simple starting method in injury severity analyses. Even though, this study attempted to capture factors that increase the likelihood of commercial drivers being involved in injury and fatal ROR crashes, the drawn inferences will be erroneous because heterogeneity was ignored.</td>
</tr>
</tbody>
</table>
Table 1.1: Summary of previous studies analyzing injury severities of ROR crashes (Continued)

<table>
<thead>
<tr>
<th>Study</th>
<th>Data used</th>
<th>Methodological approach</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Peng et al., 2012)</td>
<td>Six-year of crash data for the period from 2003 to 2008 were utilized in the analyses conducted in this study regarding injury severity. These data were obtained from Texas Department of Transportation (TxDOT) crash records information system (CRIS).</td>
<td>A multinomial logit model was used for crash severity analysis.</td>
<td>This study primarily focused on the effect of roadside characteristics on single-vehicle ROR crashes on rural two-lane roadways. Once again, heterogeneity in crash data was ignored in the analyses. Therefore, the estimation results are suspected to be inaccurate.</td>
</tr>
<tr>
<td>(Roy and Dissanayake, 2011)</td>
<td>Crash data pertaining to ROR and non-ROR crashes for the period between 1999 and 2008 in the state of Kansas were obtained from the Kansas Department of Transportation (KDOT). Injury severity was used as a binary variable.</td>
<td>The Bayesian statistical approach was utilized.</td>
<td>Despite that the Bayesian statistical method relaxes some assumptions that are required in regression models, misleading results are more likely to be obtained because Bayesian inferences require skills to translate subjective prior beliefs into a mathematically formulated prior. Moreover, issues inherently associated with crash data such as heterogeneity did not account for. Lastly, ROR crashes, in general, were considered in the sense that there was no specific disaggregation.</td>
</tr>
<tr>
<td>(Montella and Pernetti, 2010)</td>
<td>Crash data pertaining to the ROR crashes for the period between 2001 and 2005 were utilized. Injury severity was grouped into six outcomes.</td>
<td>Chi-squared test with Yates’ correction was used for analysis.</td>
<td>This study has focused on applying the chi-squared test for comparison between injury severity groups and factors affecting those groups. Still, discrete choice models, the appropriate methods for injury severity analyses, were not used. Moreover, heterogeneity in crash data was not accounted for.</td>
</tr>
</tbody>
</table>
Table 1.1: Summary of previous studies analyzing injury severities of ROR crashes (Continued)

<table>
<thead>
<tr>
<th>Study</th>
<th>Data used</th>
<th>Methodological approach</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Dissanayake, 2003)</td>
<td>The crash data used were extracted from the Florida Traffic Crash Database. These data included ROR crashes involving young drivers with 15 to 20 years old that occurred between 1997 and 1998.</td>
<td>A binary logistic regression model was used.</td>
<td>Injury severity was coded into five outcomes: no injury, possible injury, non-incapacitating injury, incapacitating injury, and fatal. However, sequential binary logistic regression models were used between each pair of injury levels. The authors instead could use other discrete choice models that account for heterogeneity that arises in crash data. Also, they just considered young drivers without any comparison with other age groups.</td>
</tr>
<tr>
<td>(Lee and Mannering, 2002)</td>
<td>Crash data used were concerning ROR crashes occurred between 1994 and 1996 in the state of Washington.</td>
<td>A nested logit model was employed for injury severity.</td>
<td>This study examined injury severity of ROR crashes. However, heterogeneity in crash data was ignored. Furthermore, ROR crashes, in general, were considered without disaggregation.</td>
</tr>
</tbody>
</table>
References


Montella, A., Imbriani, L.L., Mauriello, F., Pernetti, M., 2015. Factors Contributing to


Ye, F., Lord, D., 2011. Investigation of Effects of Underreporting Crash Data on Three Commonly Used Traffic Crash Severity Models: Multinomial Logit, Ordered Probit, and Mixed Logit. Transportation Research Record: Journal of the

AN EMPIRICAL ANALYSIS OF RUN-OFF-ROAD INJURY SEVERITY CRASHES INVOLVING LARGE TRUCKS

By: Nabeel Saleem Saad Al-Bdairi and Salvador Hernandez, Ph.D.

Journal: Accident Analysis and Prevention
Address: https://doi.org/10.1016/j.aap.2017.02.024
Issue: 102 (2017) pg. 93-100
Chapter 2: An empirical analysis of run-off-road injury severity crashes involving large trucks

BY: Nabeel Saleem Saad Al-Bdairi
Graduate Student, School of Civil and Construction Engineering
Oregon State University, 233 Owen Hall, Corvallis, OR 97331-3212
Email: albdairn@oregonstate.edu

Salvador Hernandez, Ph.D.
Assistant Professor, School of Civil and Construction Engineering
Oregon State University, 309 Owen Hall, Corvallis, OR 97331-3212
Tel: 541-737-4740
sal.hernandez@oregonstate.edu
Abstract

In recent years, there has been an increasing interest in understanding the contributory factors to run-off-road (ROR) crashes in the US, especially those where large trucks are involved. Although there have been several efforts to understand large-truck crashes, the relationship between crash factors, crash severity, and ROR crashes is not clearly understood. The intent of this research is to develop statistical models that provide additional insight into the effects that various contributory factors related to the person (driver), vehicle, crash, roadway, and environment have on ROR injury severity. An ordered random parameter probit was estimated to predict the likelihood of three injury severity categories using Oregon crash data: severe, minor, and no injury. The modeling approach accounts for unobserved heterogeneity (i.e., unobserved factors). The results showed that five parameter estimates were found to be random and normally distributed, and varied across ROR crash observations. These were factors related to crashes that occurred between January and April, raised median type, loss of control of a vehicle, the indicator variable of speed not involved, and the indicator variable of two vehicles or more involved in the crashes. In contrast, eight variables were found to be fixed across ROR observations. Looking more closely at the fixed parameter results, large-truck drivers who are not licensed in Oregon have a higher probability of experiencing no injury ROR crash outcomes, and human related factor, fatigue, increases the probability of minor injury. The modeling framework presented in this work offers a flexible methodology to analyze ROR crashes involving large trucks while accounting for unobserved heterogeneity. This information can aid safety planners and the trucking industry in identifying appropriate countermeasures to help mitigate the number and severity of large-truck ROR crashes.

Keywords: Large-truck safety, run-off-road crashes, ordered random parameter probit model, injury severity, unobserved heterogeneity
2.1 Introduction

In recent years, there has been an increasing interest in understanding the contributory factors to run-off-road (ROR) crashes in the US, especially those where large trucks are involved (Davis et al., 2006; Lee and Mannering, 2002; McLaughlin et al., 2009; Peng and Boyle, 2012; Roy and Dissanayake, 2011). One reason for this is that in 2010, approximately 57% of all fatal crashes were ROR crashes, whereas nonfatal crashes accounted for 16% (Blincoe et al., 2015). Accordingly, those crashes led to roughly $64 billion in economic costs and $298 billion in comprehensive costs, accounting for 27% of all economic costs and 36% of all societal harm (Blincoe et al., 2015). Although, statistically, the number of large-truck-involved crashes has decreased over the past two decades, there is still a higher fatal crash involvement rate per 100 million vehicle miles traveled compared to passenger cars (1.34 versus 1.08 for the year 2014) (Federal Motor Carrier Safety Administration, 2016). As a result, several state agencies have today developed and/or have adopted mitigation programs to reduce the number and severity of these crashes. For example, in Oregon, where nearly 66% of all fatal crashes were due to ROR crashes in 2010, the Oregon Department of Transportation (ODOT) partnered with the Federal Highway Administration (FHWA) to implement appropriate and low-cost countermeasures with the goal of reducing the number of ROR fatalities by 20%. However, the implemented countermeasures focused primarily on reducing ROR crashes for passenger cars, with little focus on large trucks (gross vehicle weight rating [GVWR] greater than 10,000 pounds). With this in mind, there is a clear need for continued research into identifying and/or developing cost-effective countermeasures to reduce the number and severity of ROR crashes involving large trucks.

Various methodological models have been used in analyzing severity and frequency of crashes. The selection of an appropriate statistical model depends primarily on the nature of the crash data. In this study, the injury severity sustained by trucks’ drivers are the main interest. Therefore, only previous works that examined the crash severity will be reviewed. The current study has aimed to examine the impact of contributory factors on the run-off-the-road (ROR) crashes that involved large trucks. In terms of
the risk factors, several studies have been conducted to investigate the effect of some possible factors on ROR crash severity. In general, these factors can be summarized into three main groups, human factors, highways’ geometric design and environmental factors, and roadside factors. In terms of human factors, Davis et al. (2006) reviewed the previous works that studied the effect of speed on ROR crashes on rural two-lane highways. They collected data from Australia and Minnesota for their study and used Bayesian relative risk regression. In their study, they found that high speed was associated with higher fatality risk. McGinnis et al. (2001) used Fatality Analysis Reporting System (FARS) data for years 1975, 1980, 1985, 1990, 1996, and 1997 to investigate the contributory factors that might affect ROR crashes. They found that around half of ROR crashes occurred due to intoxicated drivers, particularly male drivers whose age between 20 and 39 years. They also found that severity of ROR crashes that involved male drivers were higher than ROR crashes that involved female drivers.

Turning to statistical approaches, several models have been used to determine the relationship between the potential contributory factors and ROR crash severity. Liu and Subramanian (2009) conducted a univariate analysis with chi-square tests and logistic regression to analyze FARS crash data for the period from 1991 to 2007. In their study, they attempted to capture the effect of various factors on ROR crashes such as roadway and environmental factors, driver characteristics, and traffic-related factors. They stated that some variables were statistically significant and affected ROR crashes such as the presence of horizontal curves on the roadway, alcohol impairment, number of lanes, inclement weather, and driver age. Roy and Dissanayake (2011) developed a Bayesian statistical approach to compare ROR with non-ROR crashes in Kansas by using crash data for crashes that occurred in the period between 1999 and 2008. They found some variables were highly associated with ROR crashes rather than non-ROR crashes. These variables included road surface condition (i.e., wet and icy surfaces), time of the day (i.e., nighttime), rural area, inclement weather conditions, horizontal curve sections, higher speed, and fatigue and drowsiness. Dissanayake (2003) conducted a study to identify the contributory factors that affect the severity of ROR crashes involving young drivers with age from 16 to 25 years. In this study,
Dissanayake obtained a crash data from Florida traffic crash database. He categorized the injury severity into five injury outcomes and then developed four sequential binary logistic regression models. He concluded that some factors were highly influencing the severity of young drivers involved in ROR crashes such as gender, lighting condition, area type, and roadway alignment.

The majority of the aforementioned studies primarily focused on analyzing the injury severity of passenger cars involved ROR crashes. However, works that study the injury severity of drivers involved in large trucks ROR crashes are sparse. Some studies in recent years have specifically studied large-truck-involved crashes from various perspectives. Some of this work has dealt with understanding the risk and human-related factors of ROR crashes caused by speed, driver characteristics, driving under the influence of alcohol and/or drug impairments, fatigue or drowsiness, roadway characteristics, vehicle characteristics, and environmental factors (Aram, 2010; Compton and Berning, 2009; LeRoy et al., 2008; McGinnis et al., 2001; Neuman et al., 2003; Peng and Boyle, 2012; NHTSA, 2012). Other studies have focused on identifying the contributory factors to large-truck-involved crashes through econometric and statistical models. In those studies, ROR crashes are represented as an indicator variable for crashes related to urban settings, rural versus urban, time of day, manner of collision, and vehicle type (Cerwick et al., 2014; Chang and Mannering, 1999; Chen and Chen, 2011; Duncan et al., 1998; Islam and Hernandez, 2013a; Islam and Hernandez, 2013b; Islam and Hernandez, 2015; Islam et al., 2014; Khorashadi et al., 2005; Lemp et al., 2011; Pahukula et al., 2015; Romo et al., 2014).

Although there have been several efforts to understand large-truck crashes, the relationship between contributory factors and severity of ROR crashes is not clearly understood. One reason for this stems from the lack of detailed crash data to capture the complex interactions of multiple factors under a single framework for ROR crashes. Therefore, the purpose of this research is to develop statistical models that provide additional insight into the effects that various contributory factors related to the person (driver), vehicle, crash, roadway, and environment have on ROR injury severity. This is done by analyzing the Oregon Statewide Crash Data System, which is an extensive database collected and maintained by ODOT. The findings of this study can provide
information that can aid safety planners and the trucking industry in identifying appropriate countermeasures to help mitigate the number and severity of large-truck ROR crashes. To the best of the authors’ knowledge, these are the first attempts at developing these types of models for ROR crashes.

The rest of the paper is organized as follows. In section 2, the crash data used in the analysis and their descriptive statistics are described. Section 3 presents details of the proposed econometric modeling framework. In section 4, estimation results along with discussions are presented. Section 5 provides conclusions and suggestions for future research.

2.2 Data Description

This study utilizes data collected from the Oregon Statewide Crash Data System provided by ODOT. The data obtained represents seven years of large-truck-involved crashes, from 2007 to 2013; large-truck-involved crashes for the seven-year period comprised 13,364 records. However, since ROR crashes are the main interest of this study, only crashes belonging to this category are considered, bringing the sample size down to roughly 2,486 observations (data filtered by ROR indicator). Each ROR observation represents the maximum level of injury severity sustained by the driver following the National Safety Council (NSC) injury severity scale, KABCO. The KABCO injury severity scale characteristically consists of five injury categories: fatality (K), incapacitating (A), non-incapacitating (B), possible injuries (C), and non-injury (O) or property damage only (PDO). For this study, any recorded incidents that showed an injury severity of “not reported” or “unknown” were rejected because the severity of those injuries could not be satisfactorily determined. As was the case with other studies (Anarkooli and Hosseinlou, 2016; Haleem and Abdel-Aty, 2010; Haleem and Gan, 2013; Pahukula et al., 2015; Quddus et al., 2002), because of low data observations for the higher injury severity outcomes, the full KABCO scale was reduced to three injury categories. These categories are severe injury (KA- fatal and incapacitating), minor injury (BC- non-incapacitating and possible injury), and no injury (O- property damage only or PDO).
Turning to the data, overall severe injury, minor injury, and no injury accounted for 2.6% (N = 65), 24.6% (N = 612), and 72.8% (N = 1809), respectively. Table 2.1 illustrates the descriptive statistics of key variables for large-truck-ROR crash severity. These variables were selected according to their statistical significance and minimal correlation.

In terms of the driver-related factors, injury statistics, shown in Table 2.1, show that 2.4% of truck drivers who were involved in ROR crashes sustained severe injury when they were not under the influence of alcohol at the time of the crash, whereas 72.9% of those drivers sustained no injury. Moreover, 75.5% of drivers whose driver license were not issued in Oregon were experienced no injury, while drivers who were sustained severe injury were 1.9%. The potential reason might be related to driver behavior, probably because their driving is more cautious since they are unfamiliar with Oregon highways. Furthermore, fatigued drivers were found to be more likely to sustain minor- and no-injury since injury statistics show that those injury categories are 43.6% and 56.4%, respectively. In addition, losing control of a vehicle constituted as highly as effect than aforementioned driver related factors because 3.5% of ROR crashes that occurred due to losing control of vehicles caused severe injury for drivers.

With regard to roadway and environmental factors, ROR crashes that occurred on horizontal curves and on dry roadway surfaces were associated with high probability of sustaining severe injury. Table 2.1 depicts that those crashes have slightly similar impact on drivers’ injury since 3.5% of ROR crashes on horizontal curves and 3.6% of ROR crashes on dry roadway surfaces caused severe injury for drivers. Moreover, it is quite interesting to note that 77.5% of ROR crashes that occurred between January and April were associated with no injury outcome, whereas minor and severe injury outcomes were 20.7% and 1.8%, respectively. One potential explanation is Oregonian drivers might be accustomed to the prevalent adverse and inclement weather conditions in the aforementioned period. For more details regarding percentage distribution of ROR crashes for each injury category (see Table 2.1).
Table 2.1: Frequency and percentage distribution of driver injury

<table>
<thead>
<tr>
<th>Variable</th>
<th>Severe injury</th>
<th>Minor injury</th>
<th>No injury</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Month of the year (1 if crash occurred between January and April, 0 otherwise)</td>
<td>13 (1.8%)</td>
<td>148 (20.7%)</td>
<td>555 (77.5%)</td>
<td>716</td>
</tr>
<tr>
<td>Median type (1 for raised median, 0 otherwise)</td>
<td>6 (1.3%)</td>
<td>73 (16.9%)</td>
<td>354 (81.8%)</td>
<td>433</td>
</tr>
<tr>
<td>Roadway surface condition (1 for dry, 0 otherwise)</td>
<td>50 (3.6%)</td>
<td>387 (27.5%)</td>
<td>969 (68.9%)</td>
<td>1406</td>
</tr>
<tr>
<td>Driver license status (1 for a license from other states or countries, 0 otherwise)</td>
<td>20 (1.9%)</td>
<td>242 (22.6%)</td>
<td>810 (75.5%)</td>
<td>1072</td>
</tr>
<tr>
<td>Vehicle maneuver just before impending crash (1 if going straight, 0 otherwise)</td>
<td>62 (3.0%)</td>
<td>578 (28.1%)</td>
<td>1417 (68.9%)</td>
<td>2057</td>
</tr>
<tr>
<td>Participant level action (1 for lost control of a vehicle, 0 otherwise)</td>
<td>32 (3.5%)</td>
<td>291 (32.2%)</td>
<td>582 (64.3%)</td>
<td>905</td>
</tr>
<tr>
<td>Participant level safety equipment use (1 if seatbelt was fastened, 0 otherwise)</td>
<td>33 (1.6%)</td>
<td>511 (25.7%)</td>
<td>1445 (72.7%)</td>
<td>1989</td>
</tr>
<tr>
<td>Crash type (1 for overturn, 0 otherwise)</td>
<td>8 (2.0%)</td>
<td>182 (45.5%)</td>
<td>210 (52.5%)</td>
<td>400</td>
</tr>
<tr>
<td>Alcohol not a factor (1 for no alcohol involved, 0 otherwise)</td>
<td>59 (2.4%)</td>
<td>605 (24.7%)</td>
<td>1789 (72.9%)</td>
<td>2453</td>
</tr>
<tr>
<td>Speed not a factor (1 for non-speed involved, 0 otherwise)</td>
<td>62 (2.5%)</td>
<td>589 (24.2%)</td>
<td>1784 (73.3%)</td>
<td>2435</td>
</tr>
<tr>
<td>Roadway characteristics (1 for horizontal curve, 0 otherwise)</td>
<td>25 (3.5%)</td>
<td>251 (35.5%)</td>
<td>431 (61.0%)</td>
<td>707</td>
</tr>
<tr>
<td>Participant level action (1 for fatigue, 0 otherwise)</td>
<td>0 (0.0%)</td>
<td>48 (43.6%)</td>
<td>62 (56.4%)</td>
<td>110</td>
</tr>
<tr>
<td>Number of vehicles involved in the crash (1 if two or more vehicles, 0 otherwise)</td>
<td>6 (1.0%)</td>
<td>64 (10.6%)</td>
<td>532 (88.4%)</td>
<td>602</td>
</tr>
</tbody>
</table>

It should be noted that some variables, such as the type of shoulder, shoulder width, lane width, and the number of lanes, were not considered in the analysis. This is because the information regarding those variables was unavailable or the crash dataset had a large proportion of missing data for the variables not considered. Furthermore, the 95% confidence level was used to gauge the statistical significance of the selected variables.
(see Results and Discussion section). Consequently, statistically insignificant variables were not considered.

### 2.3 Methodological Approach

With regard to methodological approaches, many applications of statistical modeling methods have been applied in recent years to a variety of injury severity analysis scenarios. Mannering and Bhat (2014) provide a complete review of these applications. However, most of these studies have focused primarily on crash data related to a passenger car or all traffic mixes in a single modeling framework. With regard to large-truck-involved crashes, studies on modeling injury severity analysis are sparse, and a variety of statistical modeling frameworks are used, depending on the definition of the variables of interest (e.g., how injury severity is defined) (Cerwick et al., 2014; Chen and Chen, 2011; Duncan et al., 1998; Islam and Hernandez, 2013a; Islam and Hernandez 2013b; Islam et al., 2014; Khorashadi et al., 2005; Lemp et al., 2011; Pahukula et al., 2015). In this work, three injury categories are used, as previously defined, to model injury severity of ROR crashes: severe injury, minor injury, and no injury.

With this in mind, to investigate the relationship between the injury severity of ROR crashes and the possible contributory factors, an ordered probit modeling framework is considered. Traditionally, ordered probit models have been used to model and account for the ordinal nature of injury severity data. However, it has been shown that the traditional (or fixed parameter) ordered probit model is susceptible to underreporting of crash injury severity. Thus, the fixed parameter ordered probit framework imposes restrictions on the impacts of explanatory variables that are assumed to be the same across individual injury observations, whereas in ordered random parameter probit model explanatory variables are assumed to be varied across the injury observations (Eluru et al., 2008; Eluru and Yasmin, 2015; Russo et al., 2014; Savolainen et al., 2011). These drawbacks in ordered probit model with fixed parameters can lead to inconsistent (i.e., incorrect) estimates of the effects of variables on injury severity (Abdel-Aty, 2003; Abdel-Aty and Keller, 2005; Anarkooli and Hosseinzadeh, 2016;
Haleem and Abdel-Aty, 2010; Obeng, 2011). Therefore, two approaches have been proposed to overcome the restrictions in the traditional ordered probit model (Eluru and Yasmin, 2015). These approaches are either allowing thresholds to be random or the effect of explanatory variables treated as random (Eluru and Yasmin, 2015). Some previous works have been conducted in context of traffic injury by following the first approach proposed by Eluru and Yasmin (2015) (i.e., treating thresholds as random) to relax the aforementioned restrictions in traditional ordered probit models. For instance, Eluru et al. (2008), Srinivasan (2002), and Yasmin and Eluru (2013) used mixed generalized ordered logit (MGOL) to account for unobserved heterogeneity in the effect of explanatory variables on injury severity levels in both the latent injury risk propensity function and the threshold functions.

On the other hand, Eluru et al. (2008), Eluru and Yasmin (2015), Hensher et al. (2015), Russo et al. (2014), and Savolainen et al. (2011) argue that to overcome the aforementioned drawbacks of the ordered probit model with fixed parameter, extending the ordered probit model to an ordered random parameter probit model that accounts for unobserved heterogeneity (also referring to unobserved factors), can account for the above drawbacks. In this paper, the impact of explanatory variables has been treated as random to overcome the restrictions of traditional ordered probit model. To illustrate the superiority of the ordered random parameter probit model, this model will be compared with the fixed parameter ordered probit model.

2.3.1 Ordered Random Parameter Probit Model

The ordered random parameter probit model is formulated by specifying an unobserved variable, \( z \), as a linear function of a vector of explanatory variables \( X \) and the associated vector of estimable parameters \( \beta \) (e.g., person [driver], vehicle, crash, roadway, and environment), along with an error term or a disturbance term \( \varepsilon \), which is assumed to be independently randomly distributed, with a mean of 0 and a variance of 1 (Eluru et al., 2008; Hensher et al., 2015; Washington et al., 2011). The unobserved variable, \( z \), can be represented as illustrated in Eq. (2.1):
\[ z = \beta X + \epsilon \]  \hspace{1cm} (2.1)

Then, by using Eq. (2.1) for each observation, ordinal injury data \( y \) can be defined as shown in Eq. (2.2)

\[
\begin{align*}
    y = 1 & \quad \text{if } z \leq \mu_0 \\
    y = 2 & \quad \text{if } \mu_0 < z \leq \mu_1 \\
    y = 3 & \quad \text{if } \mu_1 < z \leq \mu_2 \\
    y = \ldots & \\
    y = J & \quad \text{if } z \geq \mu_{J-1}
\end{align*}
\]  \hspace{1cm} (2.2)

where \( \mu \) equals estimable parameters or thresholds that define the ordinal injury data \( y \). In general, these thresholds are estimated jointly with model estimable parameters \( \beta \), which corresponds to integer ordering, and where \( J \) represents the highest integer ordered response (in this work, that response is no injury).

Next, to estimate the probabilities of \( J \), the specific ordered response for each ROR crash observation, the error term (\( \epsilon \)) is assumed to be normally distributed, with mean and variance equal to 0 and 1, respectively. The ordered selection probabilities are illustrated in Eq. (2.3):

\[
\begin{align*}
    P(y = 1) &= \Phi (-\beta X) \\
    P(y = 2) &= \Phi (\mu_1 - \beta X) - \Phi (-\beta X) \\
    P(y = 3) &= \Phi (\mu_2 - \beta X) - \Phi (\mu_1 - \beta X) \\
    \ldots & \\
    P(y = J) &= 1 - \Phi (\mu_{J-1} - \beta X)
\end{align*}
\]  \hspace{1cm} (2.3)

where the highest (\( y = 3 \)) represents no injury and the lowest (\( y = 1 \)) represents severe injury.
As an attempt to account for unobserved heterogeneity and to allow a variable to have various effects across the observations, an ordered probit model with random parameters is applied. The simulated maximum likelihood estimation procedure was established by Greene to use random parameters in the ordered probit models, as illustrated in Eq. (2.4) (Greene, 2007).

\[ \beta_i = \beta + u_i \]  

(2.4)

where \( u_i \) is a randomly distributed term (for example, a normally distributed term, with mean 0 and variance \( \sigma^2 \)). Estimation of the ordered random parameter probit model is accomplished through the use of the Halton sequence approach (Anastasopoulos and Mannering, 2009; Bhat, 2003; Train, 1998). In this study, 200 Halton draws were used, and several random parameter distributions were examined, such as the normal, lognormal, triangular, and uniform distributions (Anastasopoulos and Mannering, 2009). However, only the normal distribution produced statistically significant results.

To investigate the impact of a particular variable on the injury outcomes, the marginal effects are used. The marginal effects represent the change in the probability of a particular injury outcome due to one unit change in an explanatory variable while holding all other variables constant. Moreover, the marginal effects are commonly used along with an ordered probit model to help interpret the interior injury outcomes or thresholds. The marginal effects corresponding to the probability of each category can be estimated as illustrated in Eq. (2.5) (Washington et al., 2011):

\[
\frac{\partial P(y = 1)}{\partial x} = -\phi (-\beta x) \beta'
\]

\[
\frac{\partial P(y = 2)}{\partial x} = [\phi (\mu_0 - \beta x) - \phi (\mu_1 - \beta x)] \beta'
\]

\[
\frac{\partial P(y = 3)}{\partial x} = [\phi (\mu_1 - \beta x) - \phi (\mu_2 - \beta x)] \beta'
\]

\[
\frac{\partial P(y = J)}{\partial x} = -\phi (\mu_{J-1} - \beta x) \beta'
\]

(2.5)
2.4 Results and Discussion

Using simulation-based maximum likelihood and maximum likelihood methods with 200 Halton draws, an ordered probit model with random parameters was estimated. The purpose for using 200 Halton draws was to get a precise and accurate estimate of the random parameters (Bhat, 2003). With regard to the distribution of the random parameters in this analysis, the uniform, triangular, lognormal, and normal distributions were tested, and the normal distribution was the only distribution that resulted in statistically significant estimates for the random parameters. The econometric software NLOGIT 5.0 was used to analyze the data and fit the ordered probit model with fixed parameters and random parameters. The estimated results for the ordered probit model with fixed and random parameters along with the marginal effects are shown in Tables 2.2 and 2.3, respectively.

The estimated results presented in Table 2.2 illustrate that only five variables were found to be random and statistically significant (the standard deviation was statistically significant or different from zero), whereas the rest of other variables, along with the constant term, were found to be fixed variables and did not vary across the observations. The random variables were crashes that occurred between January and April, raised median type, loss of control of a vehicle, the indicator variable of speed not involved, and the indicator variable of two vehicles or more involved in the crashes. All the random variables were normally distributed. The criterion that distinguishes between the fixed- and random-parameters in a model is the statistical significance of the standard deviation corresponding to each variable. In other words, if the standard deviation corresponding to a particular variable is significant, and different from zero, that variable will be a random variable, and vice versa (Agbelie, 2014).
Table 2.2: Estimated Results for the Ordered Probit Model with Fixed and Random Parameters (Values in Parentheses Indicate the Standard Deviation of the Random Parameters Distribution)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Fixed Parameter Model</th>
<th>Random Parameter Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>1.772</td>
<td>1.863</td>
</tr>
<tr>
<td>Month of the year (1 if crash occurred between January and April, 0 otherwise)</td>
<td>0.122 1.90</td>
<td>0.161 (0.268) 2.35 (4.72)</td>
</tr>
<tr>
<td>Median type (1 for raised median, 0 otherwise)</td>
<td>0.262 3.23</td>
<td>0.361 (0.448) 3.99 (5.44)</td>
</tr>
<tr>
<td>Roadway surface condition (1 for dry, 0 otherwise)</td>
<td>-0.214 -3.56</td>
<td>-0.239 -3.79</td>
</tr>
<tr>
<td>Driver license status (1 for a license from other states or countries, 0 otherwise)</td>
<td>0.162 2.83</td>
<td>0.178 2.95</td>
</tr>
<tr>
<td>Vehicle maneuver just before impending crash (1 if going straight, 0 otherwise)</td>
<td>-0.755 -7.71</td>
<td>-0.873 -8.24</td>
</tr>
<tr>
<td>Participant level action (1 for lost control of a vehicle, 0 otherwise)</td>
<td>-0.242 -3.97</td>
<td>-0.239 (0.141) -3.91 (3.22)</td>
</tr>
<tr>
<td>Participant level safety equipment use (1 if seatbelt was fastened, 0 otherwise)</td>
<td>0.223 3.18</td>
<td>0.276 3.83</td>
</tr>
<tr>
<td>Crash type (1 for overturn, 0 otherwise)</td>
<td>-0.310 -4.45</td>
<td>-0.322 -4.25</td>
</tr>
<tr>
<td>Alcohol not a factor (1 for no alcohol involved, 0 otherwise)</td>
<td>0.592 2.75</td>
<td>0.682 3.35</td>
</tr>
<tr>
<td>Speed not a factor (1 for non-speed involved, 0 otherwise)</td>
<td>0.342 2.03</td>
<td>0.335 (0.133) 1.94 (4.58)</td>
</tr>
<tr>
<td>Roadway characteristics (1 for horizontal curve, 0 otherwise)</td>
<td>-0.207 -3.47</td>
<td>-0.232 -3.74</td>
</tr>
<tr>
<td>Participant level action (1 for fatigue, 0 otherwise)</td>
<td>-0.270 -2.16</td>
<td>-0.254 -1.76</td>
</tr>
<tr>
<td>Number of vehicles involved in the crash (1 if two or more vehicles, 0 otherwise)</td>
<td>0.436 5.45</td>
<td>1.078 (1.142) 8.23 (10.17)</td>
</tr>
<tr>
<td>Threshold ($\mu$)</td>
<td>1.477 25.77</td>
<td>1.582 24.71</td>
</tr>
<tr>
<td>Number of observations</td>
<td>2486</td>
<td>2486</td>
</tr>
<tr>
<td>Log-likelihood at zero, LL(0)</td>
<td>-1669.78</td>
<td>-1669.78</td>
</tr>
<tr>
<td>Log-likelihood at convergence, LL($\beta$)</td>
<td>-1504.41</td>
<td>-1498.32</td>
</tr>
<tr>
<td>McFadden pseudo-$\rho^2$</td>
<td>0.099</td>
<td>0.103</td>
</tr>
</tbody>
</table>
Table 2.3: Marginal Effects for the Ordered Probit Model with Random Parameters

<table>
<thead>
<tr>
<th>Variable</th>
<th>Marginal Effects</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Severe injury</td>
<td>Minor injury</td>
<td>No injury</td>
<td></td>
</tr>
<tr>
<td>Month of the year (1 if crash occurred between January and April, 0 otherwise)</td>
<td>-0.0029</td>
<td>-0.0395</td>
<td>0.0424</td>
<td></td>
</tr>
<tr>
<td>Median type (1 for raised median, 0 otherwise)</td>
<td>-0.0054</td>
<td>-0.0825</td>
<td>0.0879</td>
<td></td>
</tr>
<tr>
<td>Roadway surface condition (1 for dry, 0 otherwise)</td>
<td>0.0045</td>
<td>0.0596</td>
<td>-0.0642</td>
<td></td>
</tr>
<tr>
<td>Driver license status (1 for a license from other states or countries, 0 otherwise)</td>
<td>-0.0034</td>
<td>-0.0443</td>
<td>0.0477</td>
<td></td>
</tr>
<tr>
<td>Vehicle maneuver just before impending crash (1 if going straight, 0 otherwise)</td>
<td>0.0098</td>
<td>0.1692</td>
<td>-0.1789</td>
<td></td>
</tr>
<tr>
<td>Participant level action (1 for lost control of a vehicle, 0 otherwise)</td>
<td>0.0051</td>
<td>0.0615</td>
<td>-0.0666</td>
<td></td>
</tr>
<tr>
<td>Participant level safety equipment use (1 if seatbelt was fastened, 0 otherwise)</td>
<td>-0.0066</td>
<td>-0.0736</td>
<td>0.0803</td>
<td></td>
</tr>
<tr>
<td>Crash type (1 for overturn, 0 otherwise)</td>
<td>0.0083</td>
<td>0.0871</td>
<td>-0.0954</td>
<td></td>
</tr>
<tr>
<td>Alcohol not a factor (1 for no alcohol involved, 0 otherwise)</td>
<td>-0.0303</td>
<td>-0.2012</td>
<td>0.2315</td>
<td></td>
</tr>
<tr>
<td>Speed not a factor (1 for non-speed involved, 0 otherwise)</td>
<td>-0.0097</td>
<td>-0.0935</td>
<td>0.1032</td>
<td></td>
</tr>
<tr>
<td>Roadway characteristics (1 for horizontal curve, 0 otherwise)</td>
<td>0.0052</td>
<td>0.0607</td>
<td>-0.0658</td>
<td></td>
</tr>
<tr>
<td>Participant level action (1 for fatigue, 0 otherwise)</td>
<td>0.0066</td>
<td>0.0692</td>
<td>-0.0758</td>
<td></td>
</tr>
<tr>
<td>Number of vehicles involved in the crash (1 if two or more vehicles, 0 otherwise)</td>
<td>-0.0135</td>
<td>-0.2104</td>
<td>0.2239</td>
<td></td>
</tr>
</tbody>
</table>

To test the null hypothesis that there is no statistical difference between the ordered probit model with fixed parameters and ordered probit model with random parameters in representing the provided data, a likelihood ratio test was performed. According to Washington et al. (2011), the likelihood ratio test is written as illustrated in Eq. (2.6):
\[ \chi^2 = -2[LL_{\beta \text{Fixed}} - LL_{\beta \text{Random}}] \]  

(2.6) \[ \chi^2 = -2[-1504.41 - (-1498.32)] = 12.18 \]

Together, the chi-square statistic value of 12.18 and corresponding degrees of freedom, 5 (number of random parameters), give more than 96.0% confidence level that the ordered probit model with random parameters is superior to the ordered probit model with fixed parameters. These results indicate that the null hypothesis should be rejected. Moreover, Table 2.2 shows that log-likelihood at convergence for the ordered probit model with random parameters is significantly better than the ordered probit model with fixed parameters. Furthermore, the goodness of fit, which is provided at the bottom of Table 2.2, proves that the random parameter is statistically superior.

The marginal effects, which are illustrated in Table 2.3, provide additional insights with regard to what occurs with interior injury severity categories, their corresponding probabilities, and the magnitude of change. Table 2.3 presents further details on the probability and the sign corresponding to each interior category. With regard to the interpretation of the marginal effects for ROR crashes, such as the indicator variable representing roadway surface condition (1 for dry, 0 otherwise), the probability of no injury outcome on average is -0.0642, which means that the probability of sustaining no injury outcome where ROR crashes occur on dry roadway surfaces decreases by 0.0642, on opposed to the probabilities of severe- and minor-injury outcomes that have a positive marginal effect values as illustrated in Table 2.3.

With regard to the results found in Table 2.2, environment, roadway, human (person), vehicle, and crash-related factors highlight key findings from the estimated ordered random parameter probit model. All estimated parameters included in the model were found to be statistically significant, with plausible signs. Five parameters were found to be random, with statistically significant standard deviations for the assumed distribution, which is the normal distribution. These variables were crashes that occurred between January and April, raised median type, loss of control of a
vehicle, the indicator variable of speed not involved, and the indicator variable of two
vehicles or more involved in the crashes.

2.4.1 Environmental and Roadway-Related Factors

As shown in Table 2.2, the indicator variable of crashes that occurred between
January and April was found to be statistically significant and random, with a mean of
0.161 and a standard deviation of 0.268. This finding indicates that 27.4% of ROR
crashes involving a large truck are less than zero when these crashes occurred between
January and April, whereas 72.6% of these crashes are greater than zero. That is, 27.4% of
ROR crashes that involved large trucks are less likely to result in no injury outcome,
whereas 72.6% of these crashes increase the probability of no injury outcome. As seen
from the marginal effects on Table 2.3, the probability of no injury outcome on average
is 0.0424, which is greater than the probabilities of severe- and minor-injury outcomes.
This finding may be due to the inclement weather, which is more likely to occur
between the months of January and April in Oregon, causing more cautious driving.
The finding is consistent with previous research performed by Maze et al. (2005) and
McLaughlin et al. (2009), where it was demonstrated that environmental conditions
increase the probability of ROR crashes involving large trucks.

The indicator variable representing speed not involved was also found to be random,
with a mean of 0.335 and standard deviation of 0.133. These values suggest that 0.6% of
ROR crashes where speed was not involved are less than zero, whereas 99.4% of
these crashes are greater than zero. In other words, 0.6% of ROR crashes involving
large trucks where speed was not involved are less likely to result in no injury outcome,
whereas 99.4% of these crashes are more likely to end up with no injury outcome.
Turning to Table 2.3, the probability of no injury outcome on average is 0.1032, while
the probabilities of severe- and minor-injury outcomes are less than the probability of
no injury.

The raised median is a traffic-calming device that is implemented to reduce vehicles’
speed. However, in some cases, it correlates with high crash severity, particularly for
errant vehicles. In this study, the indicator variable representing raised median was
found to be a random parameter with a mean of 0.361 and standard deviation of 0.448. These values indicate that 21.0% of ROR crashes are less than zero when these crashes occurred on a roadway with raised median, whereas 79.0% of ROR crashes under the same conditions are greater than zero. Conversely, 21.0% of ROR crashes involving large trucks that took place on roadways with raised medians are less likely to result in no injury outcome, whereas 79.0% of these crashes increase the probability of no injury outcome. Table 2.3 shows that the probability of no injury outcome is 0.0879, which is high compared to the probabilities of severe- and minor-injury outcomes. Previous research also found that the presence of a raised median was more likely to decrease crash severity. For instance, Schultz et al. (2011) stated that installing such medians on Utah roadways reduced severe crashes by 36%. Likewise, Alluri et al. (2014) found that converting two-way left-turn lanes to raised medians on Florida’s roadways was associated with a 30% reduction in crash rates. However, the randomness of this variable suggests that, for some observations, raised medians may lead to an increased probability of severe injuries.

Horizontal curves are one of the most significant contributory factors to ROR crashes. The possibility of departing the roadway at a curved section is higher than for a straight roadway section. As seen from Table 2.2, the indicator for the horizontal curve was found to be a fixed parameter, indicating that ROR crashes occurring at horizontal curves were more likely to lead to injuries that were more serious. Table 2.3 illustrates that no injury outcome will be decreased by -0.0658 when ROR crashes involving large trucks occurs on horizontal curves. Torbic et al. (2004) state that crashes on curved roadway sections are three times more frequent than those occurring on straight roadway sections. They also found that approximately 76% of fatal crashes on curved roadway sections were single ROR crashes. Islam and Hernandez (2013b) found similar results for large-truck crashes.

### 2.4.2 Human-Related Factors

Turning to human-related factors, the indicator variable representing driver license status (1 for a license from other states or countries, 0 otherwise) was found to be
statistically significant and a fixed parameter. The marginal effects in Table 2.3 indicate that for large truck drivers from other states, the probability of experiencing no injury outcome is 0.0477, which is higher than the probabilities of severe and minor injury outcomes. The possible reasons may be due to driver familiarity with Oregon highways, and this parameter estimate may also be capturing the driving complexities related to the diverse geographical nature of the state of Oregon.

The use of safety equipment (1 if a seatbelt was fastened, 0 otherwise) was also found to be significant. This fixed parameter suggests that the probability of no injury outcome on average is 0.0803 higher (see Table 2.3), while the probabilities of severe injury and minor injury are lower. Islam and Hernandez (2013b) also found the use of seatbelts to be a fixed parameter and that it increased the probability of experiencing no injury.

The parameter estimate representing being fatigued before the crash or not was found to be statistically significant and a fixed parameter. As illustrated in Table 2.3, the probability of no injury outcome on average is reduced by -0.0758 when ROR crashes involving large trucks occurred due to fatigued drivers. This finding is consistent with (Peng and Boyle, 2012). In their study, they concluded that drowsiness and fatigue were associated with severe and fatal ROR crashes.

The indicator variable for no alcohol involved (i.e., alcohol not a factor) before the crash was found to be significant for ROR crashes and a fixed parameter. Turning to Table 2.3, the probability of no injury outcome on average is 0.2315, which is higher than the probabilities of severe- and minor-injury outcomes.

The lost control of vehicle indicator variable was found to be statistically significant and a random parameter with a mean of -0.239 and standard deviation of 0.141. These values suggest that 4.5% of ROR crashes involving large trucks where the driver loss control of the vehicle are greater than zero, while 95.5% of these crashes are less than zero. In other words, 4.5% of ROR crashes involving large trucks that occurred due to loss of control of a vehicle are more likely to result in no injury outcome, whereas 95.5% of these crashes are less likely to end up with no injury outcome. Turning to Table 2.3, this suggests that for most ROR crash occurrences—and taking into account
the randomness of this parameter—the probability of no injury outcome on average is reduced by $-0.0666$. This variable may be capturing driver complexities related to vehicle performance issues, such as a flat tire, or capturing unobserved factors related to driver inattentiveness to roadway environment, hence the randomness of this variable.

2.4.3 Vehicle and Crash-Related Factors

With respect to the influence of vehicle and crash-related factors on the probability of ROR crash occurrence, the following variables were found to statistically significant. The indicator variable of two or more vehicles involved in the ROR crashes was found to be a random parameter with a mean of 1.078 and a standard deviation of 1.142. These values give parameters of less than zero for 17.3% of ROR crashes involving multiple vehicles and greater than zero for 82.7%. Specifically, 17.3% of ROR crashes involving two or more vehicles are less likely to result in no injury outcome, whereas 82.7% of these crashes increase the probability of no injury outcome. As can be seen from Table 2.3, the probability of sustaining no injury in ROR crashes involving two or more vehicles will be increased by 0.2239, which is higher than the probabilities of severe- and minor-injury outcomes. This finding may be capturing the effect of vehicle body type in reducing the impact of injury sustained by large-truck drivers (Eluru et al., 2010). Again, given the randomness of this parameter estimate, for a small portion of the ROR crash occurrences, the opposite is true.

With regard to the crash type indicator (1 for overturning, 0 otherwise), it was found to be significant and a fixed parameter. Table 2.3 shows that the probability of no injury outcome on average is $-0.0954$, which is lower than the probabilities for severe- and minor-injury outcomes.

The indicator variable of driving straight as an evasive maneuver just before an impending crash was found to be statistically significant and fixed across observations. The marginal effects of driving straight as an evasive maneuver just before an impending crash indicate that the probability of no injury outcome on average is reduced by $-0.1789$ (see Table 2.3). This finding can be substantiated by the drivers’
expectancy that they will not experience a crash on a straight roadway and it may be related to driver behavior (e.g., driver inattentiveness, vehicle performance issues).

2.5 Conclusions and Future Research

The current study explores possible contributory factors to ROR crashes that involved large trucks in Oregon, utilizing an ordered random parameter probit modeling framework. The ordered random parameter probit model is an important approach because it provides a mechanism to account and correct for unobserved heterogeneity that can arise from factors related to the driver, vehicle, road environment, weather, variations in police reporting, temporal, and other unobserved factors not captured. The data used in this study comprised crash reports taken from the state of Oregon for the years of 2007 to 2013, to the best of our knowledge a first with respect to explicitly modeling large-truck-injury severity of ROR crashes.

The results of the analyses performed provided some interesting findings. First, five parameter estimates were found to be random and varied across the ROR crash observations. These were factors related to crashes that occurred between January and April, raised median type, loss of control of a vehicle, the indicator variable of speed not involved, and the indicator variable of two vehicles or more involved in the crashes. However, driver license status, seatbelt use, crash type, alcohol not a factor, the presence of a horizontal curve, and driver fatigue were found to be fixed parameters for the ROR crashes. Looking more closely at the results of some of the significant variables, large-truck drivers who are not licensed in Oregon have a higher probability of experiencing no injury ROR crash outcomes. Another possible explanation for this outcome could be due to driver turnover suffered by the trucking industry and its impact on driver network familiarity. Second, it was discovered that median type (i.e., raised median) increases the probability of no injury outcome in ROR crashes. Third, human-related factors such as fatigued drivers have a higher probability of severe and minor injury. These findings are important from a trucking perspective because these contributory factors can be targeted through firm intervention and continued training.
Although the research performed is exploratory in nature, the ordered random parameter probit modeling framework presented in this work offers a flexible and practical methodology to analyze ROR crashes involving large trucks and to account for unobserved heterogeneity. Furthermore, this study provides information that can aid safety planners and the trucking industry in identifying appropriate countermeasures to help mitigate the number and severity of large-truck ROR crashes. Using the same approach and comparing it with recent statistical models, with an expanded sample of ROR large-truck crashes could provide important new insights into large-truck driving behavior. For example, datasets with driver skill and other cognitive processing information, car-following dynamics, and human response can greatly improve parameter estimates as well as help improve truck-driver training for collision avoidance.

Acknowledgment

We would like to acknowledge the Oregon Department of Transportation, specifically, the Transportation Data Department, for providing the crash dataset and for helping in our understanding of the data. And we would like to also acknowledge Jason Anderson for his help.
References


Davis, G., Davuluri, S., Pei, J., 2006. Speed as a Risk Factor in Serious Run-off-Road Crashes: Bayesian Case-Control Analysis with Case Speed Uncertainty. Journal


and urban driver-injury severities in accidents involving large-trucks: An exploratory analysis. Accident Analysis and Prevention 37, 910–921.


CONTRIBUTING FACTORS TO RUN-OFF-ROAD CRASHES INVOLVING LARGE TRUCKS UNDER LIGHTED AND DARK CONDITIONS

By: Nabeel Saleem Saad Al-Bdairi, Salvador Hernandez, Ph.D., Jason C. Anderson

Journal: Journal of Transportation Engineering, Part A: Systems
Address: https://doi.org/10.1061/JTEPBS.0000104
Issue: 144 (1) (2018) pg. 1-9
Chapter 3 : Contributing Factors to Run-off-Road Crashes Involving Large Trucks under Lighted and Dark Conditions

BY: Nabeel Saleem Saad Al-Bdairi
Graduate Student, School of Civil and Construction Engineering
Oregon State University, 233 Owen Hall, Corvallis, OR 97331-3212
Email: albdairn@oregonstate.edu

Salvador Hernandez, Ph.D.
Assistant Professor, School of Civil and Construction Engineering
Oregon State University, 309 Owen Hall, Corvallis, OR 97331-3212
Tel: 541-737-4740
sal.hernandez@oregonstate.edu

Jason C. Anderson
Graduate Student, School of Civil and Construction Engineering
Oregon State University, 211 Kearney Hall, Corvallis, OR 97331-3212
Email: anderjas@oregonstate.edu
Abstract

Previous studies have examined the relationships between run-off-road (ROR) crashes and the contributing factors, however the impact of lighting conditions has been insufficiently addressed. As a result, the objective of this study was to research the effect of lighting conditions on the injury severity of ROR crashes that involve large trucks. Based on the crash data pertaining to large trucks in the state of Oregon from 2007 to 2013, two separate mixed logit models were developed to capture the contributing factors that affect injury severity in each lighting condition. The levels of injury severity sustained by truck drivers were categorized into three main categories: severe injury (fatal and incapacitating), minor injury (non-incapacitating and possible injury), and no injury. The mixed logit model was used to account for unobserved factors (i.e., unobserved heterogeneity). Estimation results indicated that there are significant differences between dark and lighted conditions and that the level of injury severity outcomes was highly influenced by several complex interactions between factors and that the effects of some factors could vary across observations. The contributing factors include driver, traffic flow, roadway geometric features, land use, and time characteristics.

Keywords: Large-truck safety, run-off-road crashes, mixed logit model, injury severity, unobserved heterogeneity
3.1 Introduction

In recent years, critical issues have been addressed through transportation research related to traffic safety, particularly to road users. In general, previous studies have examined risk factors that might be correlated with a particular type of traffic crash under various conditions; still, the effect of roadway lighting on injury severity has been insufficiently addressed. For instance, even though there is less traffic volume during the darkness hours, the injury severity and traffic fatalities at nighttime are higher, particularly when there is no lighting on the roadways. A primary reason for the difference in injury severity outcomes between dark and lighted roadways is due to inadequate visibility. Inadequate visibility jeopardizes driver safety and increases the probability of traffic crashes, as a driver’s inability to detect a hazard in dark conditions decreases the driver’s likelihood to avoid it. Driving tasks are significantly affected by visibility because most of the signs, guides, and cues that direct drivers must be visually recognized.

In 2014, around 846 of 3,424 (24.7%) fatal crashes that involved large trucks, trucks with a gross vehicle weight rating (GVWR) greater than 10,000 pounds, took place on dark roadways without street lighting in the United States (Federal Motor Carrier Safety Administration, 2016). Likewise, Hasson and Lutkevich (2002) and Lutkevich et al. (2012) reported that traffic fatalities at nighttime were three times higher than those that occurred during the day. Moreover, Hasson and Lutkevich (2002) demonstrated that more than 14,000 human lives could be saved if the number of traffic crashes were the same during the day and at night. As a result, assuming $3.0 million is the value of a human life, up to $42 billion could be saved annually with a reduction in nighttime crashes.

The statistics provided show a need to better understand the effect of specific factors on crash severity, particularly for run-off-road (ROR) crashes. Moreover, previous studies have focused on the impact of light conditions on crash injury severity as an indicator variable. Further, there is indeed an urgent need to explore the effect of
unreported or inaccurate information in crash data to develop a robust statistical inference that supports the study findings. Therefore, to overcome these drawbacks in crash data, the mixed logit model is proposed. By using the mixed logit model, the influence of some attributes that are not captured in crash data (unobserved heterogeneity), including roadway characteristics, vehicle attributes, and driver behaviors, can be accounted for by allowing the estimated parameters to vary across the sample observations.

Lastly, this study can provide state agencies, the trucking industry, and transportation safety researchers with valuable and comprehensive information pertaining to the effect of lighting conditions on the propensity of involvement in ROR crashes that involve large trucks. As a result, the study findings can be useful in proposing countermeasures and policies that might save lives and reduce societal costs. Generally speaking, examining the contributing factors of injury severity for large truck drivers in ROR crashes has been overlooked in previous studies, particularly the impact of roadway lighting conditions. With this in mind, the aim of the current study is to fill the gap in literature regarding the impact of lighting conditions on injury severity of ROR crashes involving large trucks. Moreover, the literature is extended by conducting a model separation test to test if lighting conditions should be analyzed separately.

### 3.2 Literature Review

Several studies have been conducted over the years to assess the role of roadway lighting. However, analyzing the impact of lighting conditions on injury severity of ROR crashes involving large trucks is less documented. Rodegerdts et al. (2004) stated that nighttime traffic crashes could be decreased by 50% along with approximately a 43% reduction in fatalities if roadway lighting is incorporated. Elvik (1995) conducted a meta-analysis of 37 studies regarding the impact of lights on traffic crashes. Elvik found that providing unlighted roadways with lighting was accompanied by approximately a 65% decline in fatal crashes, whereas the crashes that led to a non-
fatal injury were found to be decreased by 32%. Likewise, the no injury outcome was reduced by 15%. Elvik et al. (2009) explained the effect of roadway lighting on traffic crashes based on reviewing previous studies conducted in 13 different countries, including USA, Great Britain, Switzerland, Sweden, Australia, Denmark, Japan, Finland, Israel, Germany, Norway, Netherlands, and Singapore. They found that providing these roadways with lighting was associated with a 60% decline in fatal and severe crashes, whereas the no injury crashes were decreased by 15%. Kim et al. (2013) used a mixed logit method to analyze injury severity sustained by a driver in single vehicle crashes in California, with a specific focus on age and gender. The results indicated that likelihood of fatal injury crashes increased by 92% without proper lighting.

Wu et al. (2014) found that lighting conditions in New Mexico for single- and multi-vehicle crashes are highly significant contributing factors in injury severity outcomes. They reported that a higher level of injury severity (severe and fatal) was sustained by drivers due to crashes in dark conditions. They concluded that darkness was associated with roughly a 112.9% increase in driver fatality in multiple-vehicle crashes on rural two-lane highways. Koupaenejad (2010) used multinomial logit and ordered probit models to highlight the contributing factors in a crash that involved passenger cars and large trucks. He concluded that severe injuries were more likely to be sustained by passenger vehicle occupants when crashes occurred on roadways without lighting.

Khorashadi et al. (2005) utilized a dataset pertaining to crashes that involved large trucks that occurred between 1997 and 2000 to identify the potential risk factors that affect traffic crash severities. This dataset was obtained from the California Department of Transportation (Caltrans). In their study, a multinomial logit model was developed to quantify the impact of specific factors on crash severity outcomes. They found that the probability of involvement in fatal or severe injury crashes was increased under dark conditions.
Overall, the majority of previous works have examined the impact of lighting conditions on crash injury severity using the lighting conditions as an explanatory variable. In other words, even though previous studies have focused on addressing the effect of lighting conditions on crash injury severity, the conclusive findings of these studies cannot be used as guidance for state and trucking agencies because they embodied the impact of lighting by indicator variables (e.g., effect of light conditions on specific injury outcomes). A complex interaction between the risk factors related to light conditions can create a concern regarding the validity of study findings that were obtained by representing the light conditions as indicator variables. For ROR crashes that involve large trucks, there is a need for comprehensive research to reveal the characteristics of ROR crashes under different lighting conditions and the resulting injury outcomes to propose appropriate countermeasures. In the current study, two separate mixed logit models were developed for lighting conditions (dark vs. lighted). Lighted conditions include daylight and dark with street lighting, whereas dark conditions include dark with no street lighting. Adopting this approach is particularly useful in capturing the variations in driver behaviors in dark and lighted conditions so that the estimated parameters are statistically accurate and cannot lead to an erroneous inference.

3.3 Methodological Approach

Because the crash data used in this study had a discrete unordered nature, the emphasis focused on econometric models that do not consider the ordered data (Savolainen et al., 2011). The models that deal with unordered discrete nature include the multinomial logit model, the nested logit model, and the mixed logit model (Mannering and Bhat, 2014). Although the multinomial logit model (MNL) is commonly used, its limitations restrict its suitability in modeling the crash severity. One of the major MNL limitations is the assumption regarding the disturbance terms, which are required to be independently and identically distributed (IID). Therefore, a
violation of this assumption can lead to what is referred to as the independence from irrelevant alternatives property (IIA) (Washington et al., 2011). On the other hand, the nested logit model can overcome the IIA property in the MNL by grouping alternatives that are believed to have a correlation within a nest. Thus, the aggregation might lead to uncertainty, which might cause erroneous inferences. Hence, the mixed logit model was utilized for several reasons. First, it was used to overcome the limitations of the MNL and nested logit models. Second, it accounted for the unobserved heterogeneity by using random parameters that allowed for an explanatory variable to vary across the observations (Behnood et al., 2016). The unobserved heterogeneity stems from different sources, such as variation within variables or crash data that does not provide detailed information. Accordingly, ignoring the unobserved heterogeneity can lead to biased and inefficient estimable parameters (Mannering et al., 2016).

The mixed logit model was also used to examine the effect of potential risk factors in ROR crashes that involve large trucks. Identifying the injury severity outcome formula, which is used to compute the probability of each severity level, can help in developing a formula that pertains to the mixed logit model. Following Washington et al. (2011), the discrete injury severity outcome can be determined by the following function, as shown in Eq. (3.1):

$$T_{in} = \beta_i X_{in} + \epsilon_{in}$$ (3.1)

where $T_{in}$ is a linear function of injury severity outcome $i$ for an observation $n$, $\beta_i$ is a vector of estimable parameters for injury severity outcome $i$, $X_{in}$ represents a vector of explanatory variables (e.g., variables related to the driver, vehicle, road, and environmental conditions) for determining the discrete injury severity outcome $i$ (severe, minor, and no injury) for an observation $n$, and $\epsilon_{in}$ is an error or disturbance term. In general, $X_{in}$ differs from other terms in Eq. (3.1), specifically $\beta_i$ and $\epsilon_{in}$, because it can be easily observed by the analyst, whereas other terms are not. Moreover,
to account for unobserved heterogeneity, the vector of estimable parameters $\mathbf{\beta}_i$ in Eq. (3.1) is given by the following linear formula, as shown in Eq. (3.2) (Kim et al., 2013):

$$
\mathbf{\beta}_i = m_i + M s_n + \Gamma_i \eta_{ni}
$$

(3.2)

where $m_i$ represents the fixed parameters, which are constant across all observations, $s_n$ is a matrix of factors that might cause an unobserved heterogeneity, while $M$ is a matrix of the heterogeneous variables. The third term in Eq. (3.2) represents the randomness in the equation. $\Gamma_i$ is a triangular matrix, which is used to estimate the correlation of the estimable parameters, whereas $\eta_{ni}$ is the vector for uncorrelated random variables (Kim et al., 2013). Generally, different distributional assumptions can be considered in the estimation of the random parameters. The common distributions are normal, triangular, uniform, and lognormal. To identify the random parameters in the provided data, all aforementioned distributions should be tested. Kim et al. (2013) and Mannering et al. (2016) demonstrated the criterion that could be followed to discern between fixed- and random-parameters. They asserted that a standard deviation of an explanatory variable is the key to decide whether that explanatory variable is random or fixed across the observations. If the standard deviation corresponding to the explanatory variable is not statistically significant (not different from zero), that variable will be fixed and will not vary across the observations, whereas when the standard deviation is statistically significant (different from zero), the variable is random and varies across the observations.

To determine the probability of injury outcome $i$ for observation $n$, Eq. (3.3) is utilized (Milton et al., 2008):

$$
P_{ni} = \int \frac{EXP[\mathbf{\beta}_i X_{in}]}{\sum_i EXP[\mathbf{\beta}_i X_{in}]} f(\mathbf{\beta}|\varphi) \, d\mathbf{\beta}
$$

(3.3)
where \( f(\beta | \varphi) \) is the density function of \( \beta \), whereas all other terms are defined in Eq. (3.1) and Eq. (3.2), respectively. The maximum simulated likelihood estimation (MSLE) method is typically used to estimate the mixed logit model by using Halton draws.

To identify the effect that a unit change in \( X_{ijk} \) has on the probability for crash \( i \) to result in outcome \( j \) (denoted by \( P_{ij} \)), marginal effects were calculated. The marginal effect formula used is described in Eq. (3.4) (Washington et al., 2011):

\[
M_{X_{ijk}}^p = P_{ij}(\text{given } X_{ijk} = 1) - P_{ij}(\text{given } X_{ijk} = 0)
\]

For indicator variables, the marginal effects are computed as the difference in the estimated probabilities when the indicator variables change from zero to one.

It should be mentioned that in the current study, all distributional forms were examined; however, only the normal distribution was found to be statistically significant. Moreover, two-hundred Halton draws were used to be consistent with other researchers who demonstrated that this number of draws can provide an accurate estimate regarding random parameters (Anastasopoulos and Mannering, 2009; Behnoood and Mannering, 2015; Bhat, 2003).

### 3.4 Empirical Setting

In this study, the dataset pertaining to ROR crashes that involved at least one truck in the state of Oregon was used. The datasets included police reports for the seven-year period of crashes that occurred between 2007 and 2013 and maintained by the Oregon Department of Transportation (ODOT). Crash data was filtered by vehicle type and type of crash (i.e., ROR) to include only drivers of large trucks involved in ROR crashes. The result was a dataset with 2,486 ROR crashes that involved at least one large truck. It should be mentioned that each ROR observation represents the maximum
level of injury severity sustained by a driver. The dataset includes a variable for injury severity, which is categorized into five injury levels which are: Fatal injury outcome if an involved individual dies within 30 days due to an crash (K), an incapacitating injury outcome (A), a non-incapacitating injury (B), possible injuries (C), and (O) for non-injury or property-damage only (PDO); also referred to as the KABCO injury severity scale. In the current study, the total observations that corresponded to fatal and incapacitating injuries were low. Accordingly, the injury outcomes were categorized into three main groups: severe injury (KA), minor injury (BC), and no injury (O). Several researchers have used a similar classification for injury outcomes, including (Al-Bdairi and Hernandez, 2017; Chang and Chien, 2013; Eluru et al., 2012; Pahukula et al., 2015; Wu et al., 2014). After applying the new classification, the dataset included severe injuries with 65 observations (2.6%), minor injuries with 612 observations (24.6%), and no injury with 1,809 observations (72.8%).

The literature regarding roadway infrastructures including roadway lighting is sparse. Previous studies have focused on investigating the impact of various geometric and infrastructural factors such as roadway and shoulders widths, roadway curvatures, and roadway pavement conditions on roadway safety. However, the influence of different lighting conditions is overlooked. Therefore, examining the effect of lighting conditions on the probability of involvement in ROR crashes, particularly for large trucks, and the resulting injury severity was the primary interest of this study. That being said, two different lighting conditions were considered: lighted conditions (daylight and dark with street lighting) and dark conditions (dark with no street lighting). Crashes occurred in dusk or dawn conditions were excluded. Hence, the dataset that pertained to the ROR crashes was further separated into two datasets based on whether a crash occurred in lighted or dark conditions. The dataset revealed that 634 out of 2,486 crashes occurred in dark conditions, which accounts for approximately 25.5% of the total crashes. In contrast, 1,852 out of 2,486 crashes (74.5%) occurred in lighted conditions. Tables 3.1 and 3.2 illustrate the frequency and percentages for the
selected factors that were found to be associated with ROR crashes in dark and lighted conditions for each injury level.

Table 3.1: Frequency Distribution of the Selected Variables under Dark Conditions

<table>
<thead>
<tr>
<th>Variable</th>
<th>Severe injury</th>
<th>Minor injury</th>
<th>No injury</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Injury severity</td>
<td>12 (1.9%)</td>
<td>157 (24.8%)</td>
<td>465 (73.3%)</td>
<td>634</td>
</tr>
<tr>
<td>Driver safety seatbelt (1 for not used, 0 otherwise)</td>
<td>2 (10.5%)</td>
<td>12 (63.2%)</td>
<td>5 (26.3%)</td>
<td>19</td>
</tr>
<tr>
<td>Driver was fatigued (1 for yes, 0 otherwise)</td>
<td>0 (0.0%)</td>
<td>30 (53.6%)</td>
<td>26 (46.4%)</td>
<td>56</td>
</tr>
<tr>
<td>Crash type (1 for overturn, 0 otherwise)</td>
<td>1 (0.9%)</td>
<td>44 (39.6%)</td>
<td>66 (59.5%)</td>
<td>111</td>
</tr>
<tr>
<td>Roadway characteristics (1 for horizontal curve, 0 otherwise)</td>
<td>6 (3.0%)</td>
<td>75 (37.5%)</td>
<td>119 (59.5%)</td>
<td>200</td>
</tr>
<tr>
<td>Crash hour (1 if the crash occurred between 4:00 a.m. and 6:00 a.m., 0 otherwise)</td>
<td>2 (1.1%)</td>
<td>55 (30.4%)</td>
<td>124 (68.5%)</td>
<td>181</td>
</tr>
<tr>
<td>Roadway characteristics (1 for vertical curve, 0 otherwise)</td>
<td>1 (0.8%)</td>
<td>29 (24.4%)</td>
<td>89 (74.8%)</td>
<td>119</td>
</tr>
<tr>
<td>Months (1 if crash occurred between September and December, 0 otherwise)</td>
<td>5 (1.9%)</td>
<td>75 (28.0%)</td>
<td>188 (70.1%)</td>
<td>268</td>
</tr>
<tr>
<td>Sobriety indicator (1 for sober at time of collision, 0 otherwise)</td>
<td>10 (1.6%)</td>
<td>153 (24.6%)</td>
<td>459 (73.8%)</td>
<td>622</td>
</tr>
<tr>
<td>Exceeding speed limit (1 for no, 0 otherwise)</td>
<td>11 (1.8%)</td>
<td>146 (23.8%)</td>
<td>457 (74.4%)</td>
<td>614</td>
</tr>
<tr>
<td>Median type (1 for raised median, 0 otherwise)</td>
<td>1 (0.7%)</td>
<td>23 (15.4%)</td>
<td>125 (83.9%)</td>
<td>149</td>
</tr>
<tr>
<td>Roadway surface condition (1 for dry, 0 otherwise)</td>
<td>10 (3.9%)</td>
<td>84 (32.4%)</td>
<td>165 (63.7%)</td>
<td>259</td>
</tr>
<tr>
<td>Losing control of vehicle (1 for yes, 0 otherwise)</td>
<td>5 (1.8%)</td>
<td>72 (26.1%)</td>
<td>199 (72.1%)</td>
<td>276</td>
</tr>
<tr>
<td>Crash type (1 for colliding with a fix object, 0 otherwise)</td>
<td>11 (2.5%)</td>
<td>104 (23.7%)</td>
<td>324 (73.8%)</td>
<td>439</td>
</tr>
</tbody>
</table>
Table 3.2: Frequency Distribution of the Selected Variables under Lighted Conditions

<table>
<thead>
<tr>
<th>Variable</th>
<th>Severe injury</th>
<th>Minor injury</th>
<th>No injury</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Injury severity</td>
<td>53 (2.9%)</td>
<td>455 (24.5%)</td>
<td>1344 (72.6%)</td>
<td>1852</td>
</tr>
<tr>
<td>Driver safety seatbelt (1 if not used, 0 otherwise)</td>
<td>15 (15.6%)</td>
<td>44 (45.8%)</td>
<td>37 (38.6%)</td>
<td>96</td>
</tr>
<tr>
<td>Age (1 if more than 65 years, 0 otherwise)</td>
<td>7 (7.0%)</td>
<td>16 (16.0%)</td>
<td>77 (77.0%)</td>
<td>100</td>
</tr>
<tr>
<td>Driver was fatigued (1 for yes, 0 otherwise)</td>
<td>0 (0.0%)</td>
<td>18 (33.3%)</td>
<td>36 (66.7%)</td>
<td>54</td>
</tr>
<tr>
<td>Crash type (1 for overturn, 0 otherwise)</td>
<td>7 (2.4%)</td>
<td>138 (47.8%)</td>
<td>144 (49.8%)</td>
<td>289</td>
</tr>
<tr>
<td>Driver sobriety (1 if sober, 0 otherwise)</td>
<td>49 (2.7%)</td>
<td>452 (24.7%)</td>
<td>1330 (72.6%)</td>
<td>1831</td>
</tr>
<tr>
<td>Roadway characteristics (1 for horizontal curve, 0 otherwise)</td>
<td>19 (3.7%)</td>
<td>176 (34.7%)</td>
<td>312 (61.6%)</td>
<td>507</td>
</tr>
<tr>
<td>Number of vehicles involved in the crash (1 if two cars, 0 otherwise)</td>
<td>3 (0.8%)</td>
<td>40 (10.4%)</td>
<td>342 (88.8%)</td>
<td>385</td>
</tr>
<tr>
<td>Driver residency (1 if non-Oregon resident, 0 otherwise)</td>
<td>17 (2.2%)</td>
<td>174 (22.7%)</td>
<td>576 (75.1%)</td>
<td>767</td>
</tr>
<tr>
<td>Vehicle maneuver before the crash (1 if going straight, 0 otherwise)</td>
<td>51 (3.5%)</td>
<td>422 (28.9%)</td>
<td>986 (67.6%)</td>
<td>1459</td>
</tr>
<tr>
<td>Vehicle maneuver before the crash (1 if turning right, 0 otherwise)</td>
<td>1 (0.5%)</td>
<td>10 (4.8%)</td>
<td>198 (94.7%)</td>
<td>209</td>
</tr>
<tr>
<td>Driver was ill (1 if yes, 0 otherwise)</td>
<td>6 (12.8%)</td>
<td>27 (57.4%)</td>
<td>14 (29.8%)</td>
<td>47</td>
</tr>
<tr>
<td>Area type (1 for rural, 0 otherwise)</td>
<td>44 (3.5%)</td>
<td>376 (29.7%)</td>
<td>846 (66.8%)</td>
<td>1266</td>
</tr>
<tr>
<td>Losing control of vehicle (1 for yes, 0 otherwise)</td>
<td>27 (4.3%)</td>
<td>219 (34.8%)</td>
<td>383 (60.9%)</td>
<td>629</td>
</tr>
</tbody>
</table>

3.4.1 Model Separation

To investigate the effect of attributes of ROR crashes that involve large trucks under different lighting conditions, separate mixed logit models were developed to capture
the effect of these attributes on each lighting condition. Two tests were performed to test the null hypothesis, which assumed there was no statistical difference between the separate models and a holistic model that combined them using indicator variables for lighted and dark conditions. In other words, the null hypothesis assumed that developing a statistical model for each lighting condition was not a suitable approach if the holistic model could accurately estimate the effect of the proposed attributes. The first test performed to determine whether to accept or reject the null hypothesis is a log-likelihood ratio test. The log-likelihood ratio test proposed by Washington et al. (2011) has been used by several researchers, including (Anarkooli and Hosseinzadeh, 2016; Islam et al., 2014; Pahukula et al., 2015). The log-likelihood ratio test can be illustrated in Eq. (3.5):

$$\chi^2 = -2 \left[ LL_{\text{Full}}(\beta_{\text{Full}}) - \sum_{j=1}^{J} LL_j(\beta_j) \right]$$  \hspace{1cm} (3.5)

where $LL_{\text{Full}}(\beta_{\text{Full}})$ is the log-likelihood at convergence for the holistic model, and it is equal to $-1481.10$, whereas the $\sum_{j=1}^{J} LL_j(\beta_j)$ represents the log-likelihood for the separate models that were developed. As mentioned, two separate mixed logit models were developed: one for lighted conditions with a log-likelihood at convergence equal to $-1073.05$, and the other model was developed for dark conditions with a log-likelihood at convergence equal to $-348.28$. Applying Eq. (3.5) for the known log-likelihood values yielded a chi-square ($\chi^2$) statistic of $119.54$. Then, to determine the confidence level of the null hypothesis, the degrees of freedom that correspond to the chi-square statistic were determined. The degree of freedom was 13 (the summation of estimated parameters in both dark and lighted conditions models minus the number of estimated parameters in the full model or aggregate model). Therefore, a chi-square statistic of $119.54$ with 13 degrees of freedom resulted in a $99.99\%$ confidence level.
Accordingly, the null hypothesis was rejected, and the parameters of the separate models were statistically different.

The second log-likelihood test was conducted to justify using two separate mixed logit models to examine the effect of lighting conditions on ROR crashes that involved large trucks rather than a holistic model that captures the effect of lighting conditions using indicator variables. This test is referred to as a parameter transferability test. According to Washington et al. (2011), the parameter transferability test is presented in Eq. (3.6):

\[ \chi^2 = -2[LL(\beta_{ba}) - LL(\beta_a)] \]  

(3.6)

where \( LL(\beta_{ba}) \) is the log-likelihood at convergence for model \( a \) using the converged parameters from model \( b \) (using \( b \)'s data) on lighting condition \( a \)'s data (constraining the parameters to be estimated \( b \)'s parameters), whereas \( LL(\beta_a) \) is the log-likelihood at the convergence of the model using \( a \)'s data (without constraining the parameters). The degrees of freedom corresponding to chi-square statistic \( \chi^2 \) is the number of estimated parameters in \( (\beta_{ba}) \). Therefore, restrict the estimated parameters in the lighted condition model to be the dark condition estimated parameters, and vice versa, then apply Eq. (3.6). By doing so, the values of the chi-square statistics for both cases are 1036 and 590, respectively. These values, with corresponding degrees of freedom (herein are 6 for both cases), indicate with over 99.99% confidence that the lighting conditions need to be modeled separately. Hence, developing two separate models for lighting conditions is justified, as the estimated parameters are statistically different by lighting condition.
3.5 Estimation Results

Several distributions regarding the estimation of the random parameters were considered, including normal, lognormal, triangular, and uniform; however, only the normal distribution was found to yield statistically significant results. If the standard deviation of an estimated parameter was statistically significant for the proposed distribution, the parameter is considered random and varies across observations. In contrast, statistically insignificant standard deviations (not different from 0) for a particular parameter indicates the parameter is homogenous across observations. Tables 3.3 and 3.4 show that the values of log-likelihood at convergence for both lighting conditions are statistically superior to the log-likelihood at zero, therefore indicating a better fit model.

Six parameters were found to affect injury severity of large truck drivers involved in ROR crashes in both models. The indicator variable of not wearing a seatbelt was found to be associated with severe injury in both light condition models. As shown in Table 3.3, not wearing a seatbelt increases the probability of being involved in a severe injury by 0.0078 when the crash occurred in a dark condition, while the same variable leads to an increase in the probability of a severe injury by 0.0057 when the crash occurred in lighted conditions (see Table 3.4). This finding illustrates the role of seatbelts in saving drivers’ lives and mitigating severe injuries, therefore showing a need for efforts that encourage large truck drivers to wear seatbelts. Such efforts can be undertaken by State Departments of Transportation (e.g., larger penalties for unbelted drivers), driver training programs, and the trucking industry as a whole.
Table 3.3: Mixed Logit Estimation Results for Dark Conditions (Values in Parentheses Indicate the Standard Deviation of the Random Parameters Distribution)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Parameter estimate</th>
<th>t-Stat.</th>
<th>Marginal effects</th>
<th>Severe injury</th>
<th>Minor injury</th>
<th>No injury</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Severe injury</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>-1.090</td>
<td>-2.53</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Driver safety seatbelt (1 for not used, 0 otherwise)</td>
<td>2.371</td>
<td>2.75</td>
<td>0.0078</td>
<td>-0.0022</td>
<td>-0.0056</td>
<td></td>
</tr>
<tr>
<td><strong>Minor injury</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Driver was fatigued (1 for yes, 0 otherwise)</td>
<td>2.085</td>
<td>4.44</td>
<td>-0.0009</td>
<td>0.0284</td>
<td>-0.0275</td>
<td></td>
</tr>
<tr>
<td>Crash type (1 for overturn, 0 otherwise)</td>
<td>1.242 (3.638)</td>
<td>1.57</td>
<td>(1.87)</td>
<td>-0.0005</td>
<td>0.0301</td>
<td>-0.0296</td>
</tr>
<tr>
<td>Roadway characteristics (1 for horizontal curve, 0 otherwise)</td>
<td>1.408</td>
<td>4.26</td>
<td>-0.0032</td>
<td>0.0628</td>
<td>-0.0596</td>
<td></td>
</tr>
<tr>
<td>Crash hour (1 if the crash occurred between 4:00 a.m. and 6:00 a.m., 0 otherwise)</td>
<td>0.719</td>
<td>2.25</td>
<td>-0.0010</td>
<td>0.0240</td>
<td>-0.0230</td>
<td></td>
</tr>
<tr>
<td>Roadway characteristics (1 for vertical curve, 0 otherwise)</td>
<td>-0.737 (3.791)</td>
<td>-0.57</td>
<td>(1.97)</td>
<td>0.0004</td>
<td>-0.0185</td>
<td>0.0181</td>
</tr>
<tr>
<td>Months (1 if crash occurred between September and December, 0 otherwise)</td>
<td>0.255</td>
<td>0.62</td>
<td>-0.0007</td>
<td>0.0359</td>
<td>-0.0352</td>
<td></td>
</tr>
<tr>
<td><strong>No injury</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sobriety indicator (1 for sober at time of collision, 0 otherwise)</td>
<td>2.099</td>
<td>3.69</td>
<td>-0.0282</td>
<td>-0.1977</td>
<td>0.2258</td>
<td></td>
</tr>
<tr>
<td>Exceeding speed limit (1 for no, 0 otherwise)</td>
<td>1.958</td>
<td>3.64</td>
<td>-0.0261</td>
<td>-0.1826</td>
<td>0.2087</td>
<td></td>
</tr>
<tr>
<td>Median type (1 for raised median, 0 otherwise)</td>
<td>1.040</td>
<td>2.91</td>
<td>-0.0018</td>
<td>-0.0197</td>
<td>0.0215</td>
<td></td>
</tr>
<tr>
<td>Roadway surface condition (1 for dry, 0 otherwise)</td>
<td>-0.844</td>
<td>-2.81</td>
<td>0.0069</td>
<td>0.0400</td>
<td>-0.0469</td>
<td></td>
</tr>
<tr>
<td>Losing control of vehicle (1 for yes, 0 otherwise)</td>
<td>-0.662</td>
<td>-2.12</td>
<td>0.0043</td>
<td>0.0284</td>
<td>-0.0327</td>
<td></td>
</tr>
<tr>
<td>Crash type (1 for colliding with a fix object, 0 otherwise)</td>
<td>-0.832</td>
<td>-1.91</td>
<td>0.0096</td>
<td>0.0616</td>
<td>-0.0712</td>
<td></td>
</tr>
</tbody>
</table>

Model statistics

| Number of observations                        | 634                |
| Log-likelihood at zero                       | -696.52            |
| Log-likelihood at convergence                | -348.28            |
| Adjusted $\rho^2$                            | 0.50               |
Table 3.4: Mixed Logit Estimation Results for Lighted Conditions (Values in Parentheses Indicate the Standard Deviation of the Random Parameters Distribution)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Parameter estimate</th>
<th>t-Stat.</th>
<th>Marginal effects</th>
<th>Severe injury</th>
<th>Minor injury</th>
<th>No injury</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Severe injury</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>-4.705</td>
<td>-3.06</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Driver safety seatbelt (1 for not used, 0 otherwise)</td>
<td>2.425</td>
<td>4.50</td>
<td>0.0057</td>
<td>-0.0027</td>
<td>-0.0030</td>
<td></td>
</tr>
<tr>
<td>Sobriety indicator (1 for sober at time of collision, 0 otherwise)</td>
<td>-3.565</td>
<td>-3.17</td>
<td>-0.0533</td>
<td>0.0300</td>
<td>0.0233</td>
<td></td>
</tr>
<tr>
<td>Age (1 if more than 65 years, 0 otherwise)</td>
<td>1.830</td>
<td>2.53</td>
<td>0.0034</td>
<td>-0.0019</td>
<td>-0.0016</td>
<td></td>
</tr>
<tr>
<td>Vehicle maneuver before the crash (1 if going straight, 0 otherwise)</td>
<td>2.892</td>
<td>2.51</td>
<td>0.0436</td>
<td>-0.0241</td>
<td>-0.0195</td>
<td></td>
</tr>
<tr>
<td><strong>Minor Injury</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>-3.110</td>
<td>-6.60</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Driver was fatigued (1 for yes, 0 otherwise)</td>
<td>1.625</td>
<td>2.29</td>
<td>-0.0003</td>
<td>0.0038</td>
<td>-0.0035</td>
<td></td>
</tr>
<tr>
<td>Crash type (1 for overturn, 0 otherwise)</td>
<td>3.188</td>
<td>5.17</td>
<td>-0.0023</td>
<td>0.0341</td>
<td>-0.0318</td>
<td></td>
</tr>
<tr>
<td>Vehicle maneuver before the crash (1 if turning right, 0 otherwise)</td>
<td>-3.346</td>
<td>-3.28</td>
<td>0.0009</td>
<td>-0.0060</td>
<td>0.0051</td>
<td></td>
</tr>
<tr>
<td>Roadway characteristics (1 for horizontal curve, 0 otherwise)</td>
<td>0.677</td>
<td>1.42</td>
<td>0.0018</td>
<td>0.0176</td>
<td>-0.0194</td>
<td></td>
</tr>
<tr>
<td>Losing control of vehicle (1 for yes, 0 otherwise)</td>
<td>0.541</td>
<td>0.93</td>
<td>0.0032</td>
<td>0.0302</td>
<td>-0.0335</td>
<td></td>
</tr>
<tr>
<td><strong>No injury</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Driver was ill (1 yes, 0 otherwise)</td>
<td>-6.714</td>
<td>-3.97</td>
<td>0.0020</td>
<td>0.0084</td>
<td>-0.0104</td>
<td></td>
</tr>
<tr>
<td>Area type (1 for rural, 0 otherwise)</td>
<td>0.196</td>
<td>0.33</td>
<td>0.0180</td>
<td>0.0297</td>
<td>-0.0477</td>
<td></td>
</tr>
<tr>
<td>Number of vehicles involved in the crash (1 if two vehicles, 0 otherwise)</td>
<td>3.979</td>
<td>2.10</td>
<td>0.0001</td>
<td>-0.0009</td>
<td>0.0008</td>
<td></td>
</tr>
<tr>
<td>Driver residency (1 if non-Oregon resident, 0 otherwise)</td>
<td>1.667</td>
<td>2.57</td>
<td>0.0017</td>
<td>-0.0074</td>
<td>0.0057</td>
<td></td>
</tr>
<tr>
<td><strong>Model statistic</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of observations</td>
<td>1852</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log-likelihood at zero</td>
<td>-2034.63</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log-likelihood at convergence</td>
<td>-1073.05</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adjusted $\rho^2$</td>
<td>0.47</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The next parameter that significantly affects injury severity of large truck drivers in both models is the indicator variable for a horizontal curve. The indicator variable for horizontal curve was associated with minor injury. In dark conditions, the indicator variable for a horizontal curve was homogeneous across observations, whereas in the lighted condition, the estimated parameter was found to be random and normally distributed with a mean of 0.677 and a standard deviation of 3.347. Based on the mean and standard deviation values, the normal distribution curve implies that 42% of the distribution is less than 0 and 58% of the distribution is greater than 0. In other words, approximately 42% of ROR crashes that occurred on horizontal curves were less likely to result in a minor injury, while 58% of these crashes had an increase in the likelihood of a minor injury. This variation in the effect of lighting conditions on injury severity of large truck crashes that occurred on horizontal curves may be attributed to the risk-taking behaviors of drivers when they negotiate horizontal curves. In particular, drivers tend to drive slowly and cautiously when negotiating horizontal curves; but, on the other hand, it is difficult to control large trucks on horizontal curves despite the cautious driving. Moreover, visibility is highly reduced during dark conditions and drivers may not be able to see an upcoming curve. Therefore, drivers may fail to reduce their speed to safely negotiate curves, which may not be the case in lighted conditions.

### 3.5.1 Dark Conditions

Mixed logit estimation results for the dark condition model are shown in Table 3.3. Marginal effects were calculated to examine the impact of contributing factors on injury severity, where marginal effects refer to a one-unit change in a particular variable while holding all others constant. Overall, 13 parameters were found to affect injury severity of ROR crashes in dark conditions. Among these, three parameters were random parameters, including overturning crash type, the presence of a vertical curve on the roadway, and crashes that occurred between September and December.
Regarding driver-related factors, five factors were found to significantly affect injury severity in dark conditions, and all of which were homogenous across observations. These factors are the indicator variable for driver sobriety at the time of the crash, the indicator variable for not exceeding speed limit, the indicator variable for not wearing a seatbelt at the time of a crash, fatigued drivers, and losing control of a vehicle. The first two driver related factors that were found to lead to no injury are the indicator variable for driver sobriety at the time of the crash and indicator variable for not exceeding speed limit. The marginal effects in terms of no injury for the two variables were 0.2258 and 0.2087, respectively. This finding implies that there is a 0.2258 and 0.2087 increase in the probability of no injury when drivers are sober and when drivers do not exceed the speed limit, respectively. These findings are intuitive, as sober drivers can react to unexpected hazards more quickly than intoxicated drivers. This finding is in agreement with Khattak et al. (2012) that concluded that truck drivers who were intoxicated at the time of the crash were more likely to be involved in severe crashes compared to sober drivers, and emphasizes the need for more efforts devoted to reduce impaired driving. Similarly, abiding by the speed limit can protect drivers from involvement in fatal crashes because crash severity and speed are inextricably linked.

In addition, the indicator variable for not wearing a seatbelt at the time of a crash was found to significantly affect the injury severity. Marginal effects show that the probability of sustaining a severe injury increases by 0.0078 if a seatbelt is not used. This finding reveals the importance of seatbelt enforcement, particularly among truck drivers, to reduce the loss of lives and mitigate more severe injuries. This finding is consistent with Dissanayake and Kotikalapudi (2012) that found truck drivers were less likely to be involved in severe crashes when they wore a seatbelt at the time of the crash.

The seasonal effect, which is represented by the month of the year, was also found to affect the injury outcomes of ROR crashes involving large trucks. Crashes that occurred between September and December were found to be associated with minor
injuries and marginal effects indicate there is 0.0359 increase in the probability of sustaining a minor injury. A possible explanation might be linked to weather conditions between September and December in Oregon when rainy weather is predominant. Therefore, Oregonian drivers are accustomed to these circumstances and are less likely to be involved in fatal or serious injury crashes. In addition, the estimated parameter was also found to be random with a mean of 0.255 and a standard deviation of 1.690. Based on values of the mean and standard deviation, the normal distribution curve indicates that 44% of the ROR crashes that occurred between September and December were less than 0. In other words, 44% of crashes that took place in the aforementioned period were less likely to result in minor injury outcomes, whereas 56% of these crashes were more likely to lead to minor injury outcomes. This finding varies from that found by Islam and Hernandez (2013a) in the context of heterogeneity; specifically, Islam and Hernandez (2013a) found the parameter to be homogenous for large truck crashes that occurred between September and December and less likely to lead to a non-incapacitating injury. The heterogeneous effects of this estimated parameter (between Oregon and Texas) may be related to the geographical differences between these two states and shows the importance of accounting for heterogeneity during data analysis techniques—model estimates and inferences would have been inaccurate had a non-heterogeneity method not been applied.

In term of crash-related factors, two variables were found to statistically affect injury severity; namely, overturning crashes and colliding with fixed objects. The estimated parameter for the indicator variable for an overturning crash was found to be random with a mean of 1.242 and a standard deviation of 3.638. Given these estimates, the normal distribution curve indicates that 36.6% of crashes with overturning trucks were less than 0. In other words, approximately 36.6% of crashes with overturning trucks were less likely to result in minor injury outcomes, whereas 63.4% of these crashes increased the likelihood of a minor injury. The randomness in this parameter may attempt to capture the variation in driver’s experience that may help in avoiding serious crashes and also quality of a truck compartment that protects drivers from serious
crashes. Next, colliding with a fixed object was found to decrease the likelihood of no injury, and according to marginal effects, results in a 0.0712 lower probability of no injury. The impact of darkness could increase the injury severity level of crashes that occurred due to colliding with a fixed object. The possible cause of sustaining a higher level of injury rather than no injury in fixed object crashes that occurred under dark lighting conditions is the degradation in visibility that may affect injury severity of collision with fixed objects as opposed to lighted conditions.

Regarding roadway-related variables, four variables were statistically significant and affect the injury severity level of drivers involved in ROR crashes in dark conditions. These variables include dry surface conditions, raised median type, the presence of horizontal curves, and the presence of vertical curves. With the exception of the indicator variable of vertical curves, the other variables were found to have no variation across observations. Table 3.3 illustrates the probability of minor injury is lower by 0.0185 on vertical curves compared to flat roadways. A possible explanation is that drivers tend to reduce their driving speed when they negotiate a vertical curve, particularly under dark conditions, to react faster to unexpected hazards that are not as easily identified under dark conditions (e.g., animal crossing the roadway, an oncoming vehicle that is crossing the centerline, vehicles parking on the shoulder, etc.). This estimated parameter was also found to be random with a mean of -0.737 and a standard deviation of 3.791. By using mean and standard deviation values, the normal distribution curve shows that 42.3% of ROR crashes that occurred on a vertical curve under dark conditions were greater than 0. This finding means approximately 42.3% of crashes occurred on vertical curves were more likely to result in a minor injury, while 57.7% of these crashes were less likely to cause a minor injury. The randomness in this parameter may be capturing the variation in driver behaviors related to negotiating vertical curves and the geometry of those curves. For instance, the percent grade is not provided in the data and is likely to impact severity outcomes (i.e., avoiding a crash or attempting to stop on a steep grade is more difficult than on a minor grade). In addition, the direction of travel in related to the vertical curve or if the crash occurred at the
pinnacle of the curve is not provided; therefore, these results may be attempting to capture the effects of these characteristics as well.

The indicator variable for horizontal curves was found to increase the probability of minor injuries by 0.0628. This finding agrees with Islam (2015) that found the parameter representing curved section highways was fixed and less likely to cause no injury. Regarding the impact of dry surface conditions, it was found that crashes on dry surface conditions decrease the possibility of no injury by 0.0469. Increased driving speed might be a potential factor in reducing the likelihood of no injury on dry surfaces. On dry surfaces, drivers may assume such surfaces are safe because the skid resistance is higher; therefore, they may increase their speed and the resulting crash is less likely to result in no injury (i.e., driver loses control while driving too fast and swerves into oncoming traffic or hits a fixed object, such as a concrete barrier or tree). Lastly, the presence of a raised median was associated with no injury crashes and marginal effects show that there is a 0.0215 higher probability of no injury for ROR crashes that occurred on a roadway in which a raised median was present.

3.5.2 Lighted Conditions

The estimation results of the mixed logit model for lighted conditions are presented in Table 3.4. In this model, five variables were found to have random and normally distributed estimated parameters. These parameters correspond to the indicator variables for a rural area, non-Oregonian drivers, the indicator variable of two vehicles involved in a crash, horizontal curve, and losing control of a vehicle.

The effect of losing control of a vehicle on the injury level sustained by drivers of large trucks in lighted conditions was found to be statistically significant and random. The mean of the estimated parameter for losing control of a vehicle is 0.541 and the standard deviation is 4.877. Given these estimates, the normal distribution curve implies that 45.6% of ROR crashes that occurred in lighted conditions due to losing control of a vehicle were less than 0. To illustrate, approximately 45.6% of ROR
crashes that occurred due to losing control of a vehicle crashes decreased the likelihood of a minor injury. In contrast, 54.4% of these crashes were more likely to cause a minor injury. A possible explanation for the heterogeneous effects of this variable could be related to driver behavior, such as the ability to regain control to avoid a more serious crash. On the other hand, a proportion of drivers may be less experienced in driving a large truck are unable to regain control, therefore increasing the chance of sustaining severe injury.

As shown in Table 3.4, some driver related factors were found to be statistically significant and drastically affect the injury severity of truck drivers in lighted conditions. These variables include the indicator variables of an older driver (more than 65 years), driver illness, and non-Oregonian drivers. The probability of large truck drivers older than 65 years to sustain a severe injury increases by 0.0034. One potential reason is that older drivers are characterized by particular health problems, such as vision problems, cognitive functioning, and physical changes that compromise their driving abilities. This finding could motivate state agencies to develop campaigns to suggest that older drivers drive only under good weather conditions. Also, restrictions regarding renewals of driver licenses for older drivers should be implemented to protect them and other road users from being involving in a crash.

Regarding the residency of drivers, it was found that non-Oregonian drivers are more likely to be involved in no injury crashes. This parameter was also found to be random with a mean of 1.667 and the standard deviation of 3.586. Given these estimates, the normal distribution curve indicates that 32.1% of the ROR crashes that involved non-Oregonian drivers were less than 0. This finding means that 32.1% of non-Oregonian drivers were less likely to be involved in no injury crashes, while 67.9% of those drivers were more likely to be involved in no injury outcomes. Unfamiliarity with Oregon roadways and traffic laws might encourage non-Oregonian drivers to be cautious and drive slowly. As a result, severity of a crash injury is lower for this group. Further, the variation in this parameter estimate may be attempting to capture the diverse geographical nature of Oregon compared to other states. In addition, this result
further illustrates the importance of applying a method that can capture the heterogeneous effects within specific variables. As such, state agencies can better devote efforts to mitigate crashes for all drivers and not just a percentage of them.

The area type also impacts injury severity. Table 3.4 shows that the indicator variable of a rural area decreases the likelihood of no injury crashes. Furthermore, this estimated parameter was random with a mean of 0.196 and a standard deviation of 4.682. Based on the mean and standard deviation values, the normal distribution curve shows that 48.3% of crashes that occurred in a rural area were less than 0. Specifically, 48.3% of ROR crashes that took place in a lighted rural area were less likely to cause no injury, whereas 51.7% of these crashes were more likely to cause no injury. One possible reason that substantiates the heterogeneous effect of this estimated parameter could be related to the characteristics of the rural area that is characterized by a relatively higher speed compared to urban area. Accordingly, ROR crashes in a lighted rural area would result in a higher level of injury severity. Moreover, drivers may be inclined to drive without safety equipment (i.e., seatbelt) on rural areas, as opposed to urban areas, due to less law enforcement presence in rural areas. On the other hand, some drivers tend to drive cautiously and carefully on rural areas, particularly in Oregon, to avoid colliding with crossing animals that are predominant in such locations.

Lastly, the estimated parameter for the indicator variable of two vehicles involved in ROR crashes was found to be a random with a mean of 3.979 and a standard deviation of 4.501. Given these estimates, the normal distribution curve implies that 18.8% of ROR crashes involving multiple vehicles in lighted condition were less than 0. That is to say, 18.8% of ROR crashes involving multiple vehicles in lighted conditions decrease the likelihood of no injury crashes, whereas 81.2% of these crashes are more likely to cause no injury. The randomness in this estimated parameter may be capturing the effect of vehicle body type in reducing the impact of injuries sustained by drivers of large trucks (e.g., a 2-axle truck compared to a tractor trailer). This finding is in line with Islam and Hernandez (2013b) that found the parameter representing the
number of vehicles involved in a crash was random and more likely to lead to no injury severity. These findings suggest that regardless of geographic region (e.g., Oregon, Texas, etc.), multi-vehicle crashes involving large trucks have heterogeneous effects on injury severity. That is, there are several characteristics that come into the fold when multiple vehicles are involved in a crash (i.e., crash locations, lighting conditions, vehicle preventing technology) and injury severity can be strongly influenced by these crash-specific characteristics.

### 3.6 Summary and Conclusions

Despite the large number of studies that have been conducted to address the relationships between ROR crashes and the contributing factors, the impact of lighting conditions has been insufficiently addressed. Thus, the objective of this study was to research the effect of lighting conditions on the injury severity of ROR crashes that involve large trucks. Based on the crash data pertaining to large trucks in the state of Oregon from 2007 to 2013, two separate mixed logit models were developed to capture the contributing factors that affect injury severity in each lighting condition. The mixed logit model was used to account for unobserved heterogeneity. Log-likelihood ratio tests were performed to verify the validity of using separate mixed logit models rather than one holistic model that represents both lighting conditions (i.e., lighted and dark conditions) by indicator variables. The results of the log-likelihood ratio tests revealed that using separate mixed logit models was justified. Two different lighting conditions were considered: lighted conditions (daylight and dark with street lighting) and dark conditions (dark with no street lighting).

Regarding the contributing factors that affect the injury severity of drivers of large trucks involved in ROR crashes, the estimation results revealed that there are significant differences between dark and lighted conditions. Moreover, some variables were found to affect the injury severity regardless of the lighting conditions. These variables are the indicator variable of not wearing a seatbelt, the indicator variable of a
horizontal curve, the indicator variable of fatigued drivers, the indicator variable of overturning crashes, the indicator variable of losing control of a vehicle, and the indicator variable of driver sobriety; however, their impacts on injury severity were varied.

For the dark condition model, crashes that occurred in the early morning (between 4:00 a.m. and 6:00 a.m.) are more likely to lead to minor injury crashes. The estimation results also showed that not drinking alcohol and not speeding were associated with no injury crashes. These findings reveal the importance of abiding by traffic laws regarding alcohol consumption and speed limits. Among the interesting findings in the lighted condition model, older drivers (more than 65 years) were more likely to sustain severe injuries when involved in ROR crashes. This finding is plausible because older drivers are more vulnerable to sustaining fatal or severe injuries due to health problems associated with age. Regarding lighted conditions, if large trucks are involved in ROR crashes in rural areas, the drivers are less likely to sustain no injury because these areas are usually characterized by higher speeds.

The study findings can provide insight for safety researchers and traffic agencies to identify the contributing factors and the possible causes of ROR crashes that involve large trucks as well as how these factors differ based on lighting conditions. By addressing these factors, potential countermeasures could be proposed to potentially mitigate the number and the severity of ROR crashes. In particular, since the primary interest of the current paper was to study the effect of lighting conditions on the injury severity of ROR crashes involving large trucks, it was found that the installation of roadside lights could significantly alleviate the injury severity that results from ROR crashes.

In future work, the authors will explore the effects of disaggregating the model by setting (urban or rural) and by time of day. Moreover, to obtain more in-depth results regarding the effect of lighting conditions, the data will be divided by area type and the spatial transferability of the models to other state specific datasets will be examined.
Although this research applies a formal modeling framework (i.e., mixed logit), a nonparametric analysis can be considered. While nonparametric analyses have shown good model fit, it has been shown to be the least precise due to its robustness (Greene, 2016). With such an analysis, inferences regarding the association between a discrete outcome and the covariates are broad and results in no more than a rough representation of that association (Greene, 2016). Being that nonparametric analyses do not rely on statistical distributions, these methods are often best utilized for ordinal type variables, when dealing with smaller sample sizes, or when the assumptions used for parametric methods are in question (Washington et al., 2011). In addition, in some cases (e.g., crash data) data are collected at many locations throughout a region. For such data, characteristics of different crashes may be similar if crashes are geospatially close to one another (Ott and Longnecker, 2010; Washington et al., 2011). As such, statistical methods based on the $t$ distribution result in probabilities that are different from the intended values, both in terms of confidence intervals and $p$-values (Ott and Longnecker, 2010). If this dependency exists, a more advanced analysis is required (Anselin, 1988; Greene, 2012). If the data being analyzed meets these requirements, an alternate model estimation approach may want to be explored.

**Acknowledgment**

We would like to acknowledge the Oregon Department of Transportation, specifically the Transportation Data Department, for providing the crash dataset and for helping in our understanding of the data. Findings of this study do not necessarily reflect the views of the Oregon Department of Transportation.
References


Pahukula, J., Hernandez, S., Unnikrishnan, A., 2015. A time of day analysis of crashes involving large trucks in urban areas. Accident Analysis and Prevention 75, 155–


Chapter 4: Urban and rural run-off-road (ROR) crashes and the injury severities of large truck drivers

BY: Nabeel Saleem Saad Al-Bdairi

Graduate Student, School of Civil and Construction Engineering
Oregon State University, 233 Owen Hall, Corvallis, OR 97331-3212
Email: albdairn@oregonstate.edu

Salvador Hernandez, Ph.D.
Assistant Professor, School of Civil and Construction Engineering
Oregon State University, 309 Owen Hall, Corvallis, OR 97331-3212
Tel: 541-737-4740
sal.hernandez@oregonstate.edu
Abstract

In spite of numerous efforts to quantitatively identify the factors contributing to the injury severity of different crash types in rural and urban settings, the distinction between rural and urban areas regarding the injury severity of run-off-road (ROR) crashes involving large trucks is still not clearly understood. As such, the objective of this study is to investigate the effect of area type (i.e., urban vs. rural) on injury severity outcomes sustained by drivers in ROR crashes involving large trucks while accounting for unobserved heterogeneity. To do this, the latent class ordered probit models with two classes are developed. The crash data pertaining to ROR crashes involving large trucks in Oregon between 2007 and 2014 were utilized. In this study, injury severity was grouped into four main ordered categories: no injury, possible injury, moderate injury (non-incapacitating), and severe injury (incapacitating and fatal). The estimation results reveal that the developed latent class ordered probit models (for urban and rural areas) are substantially distinct in terms of the contributing factors. However, seven parameters were found to be significant in both the urban and rural models, but with remarkable differences in terms of their impacts on the injury severity. The findings of this study could benefit trucking industry, transportation agencies, and safety practitioners to prevent or alleviate the injury severity of ROR crashes involving large trucks by developing appropriate and cost-effective countermeasures.

Keywords: Injury severity, unobserved heterogeneity, latent class ordered probit model, large trucks, run-off-road,
4.1 Introduction

Injury severity analyses have been extensively conducted over the years to better understand the factors that influence injuries sustained by drivers resulting from roadway crashes (Al-Bdairi et al., 2018; Al-Bdairi and Hernandez, 2017; Anarkooli and Hosseinlou, 2016; Anderson and Hernandez, 2017; Behnood and Mannering, 2017; Chang and Chien, 2013; Helai et al., 2008; Jung et al., 2012; Khorashadi et al., 2005; Kim et al., 2013; Lee and Li, 2014; Schneider et al., 2009; Wu et al., 2016, 2014; Xiong and Mannering, 2013). Among the many factors that have been found to influence driver injury severity, and of particular interest to this study, is that of land use type (e.g., rural vs. urban) (Al-Bdairi et al., 2018; Islam et al., 2014; Khorashadi et al., 2005; Lee and Mannering, 2002). Recent statistics from the National Highway Traffic Safety Administration (NHTSA) indicate that roadway crashes are disproportionately distributed between urban and rural areas. For instance, in 2015, approximately 35,092 individuals lost their lives due to 32,166 fatal crashes on U.S. roadways. Of these fatal crashes, about 15,293 occurred in rural areas, and roughly 14,414 took place in urban areas (NHTSA, 2017). In the state of Oregon, this distinction between rural and urban crashes and fatalities also holds true. In 2015, Oregon experienced 44,523 total crashes in urban areas, leading to 156 fatalities. In contrast, in the same year, the total crashes in rural areas were roughly one-fourth (10,633) of those reported in urban areas. However, the number of fatalities were about two times (254) more than those experienced on urban roads (ODOT, 2017).

The many reasons underlying this disparity include, but are not limited to: the longer emergency response times for individuals involved in a rural crash (Gonzalez et al., 2007); the medical intervention in rural areas is not as good as in urban areas (Zwerling et al., 2005); the higher speed limits and higher travel speeds in rural areas compared to urban areas; the lack of traffic law enforcement in rural areas compared to urban areas; the risky driving behavior in rural areas; the different traffic environments of rural and urban areas, such as traffic volume and roadway conditions (Nordfjærn et al., 2010); the lower use of protective devices, such as seatbelts, in rural areas (Yan et al.,...
2012); and the differences in individuals’ perceiving and estimating the risks of traffic crashes in rural and urban areas. These factors, coupled with other observed and unobserved (e.g., driving habits) influences, are the underlying causes of the disproportionate distribution of fatal crashes between urban and rural areas.

In spite of numerous efforts to quantitatively identify the factors contributing to the injury severity of different crash types in rural and urban settings, the distinction between rural and urban areas regarding the injury severity of run-off-road (ROR) crashes involving large trucks is still not clearly understood (Islam et al., 2014; Khorashadi et al., 2005; Lee and Mannering, 2002; Wu et al., 2016). This may be due to the land use type being treated as a contributing factor to such crashes, rather than testing for the effect of urban versus rural on injury severity. In the present study, ROR crashes involving large trucks are of particular interest for two reasons. First, large trucks play a vital role in the U.S. economy; for example, in 2013, large trucks (i.e., trucks weighing over 10,000 lbs.) moved roughly 55 million tons of freight valued at more than $49.3 billion (U.S. Department of Transportation/Bureau of Transportation Statistics, 2015). Unfortunately, the movement of this much freight does not come without a price in terms of roadway crashes and resulting fatalities (due in part to the operating and vehicle characteristics of large trucks). Second, ROR crashes are a nationwide problem that needs to be thoroughly investigated; in 2010, for instance, they constituted 57% of all fatal crashes and 16% of nonfatal crashes (Blincoe et al., 2015). Furthermore, identifying the factors contributing to the injury severity of ROR crashes involving large trucks in rural and urban areas is essential in terms of implementing practices that can favorably alter the impact of these factors. A better understanding of these contributing factors, coupled with the impact of land use setting, in relation to the injury severity sustained by large truck drivers (this refers to the size of the vehicle) involved in ROR crashes can provide transportation safety professionals, the trucking industry, and policy makers with valuable insights towards reducing the number of ROR crashes involving large trucks, and their injury severity,
through the selection and implementation of appropriate cost-effective countermeasures.

With this overarching goal in mind, the objective of this study is to investigate the effect of land use setting (i.e., urban vs. rural) on injury severity outcomes sustained by drivers in ROR crashes involving large trucks while accounting for unobserved heterogeneity (unobserved factors not present in the data) for both rural and urban perspectives. To achieve this objective, an econometric modeling framework is utilized, specifically, the latent class ordered probit model. A latent class approach provides a modeling framework that can account for possible unobserved heterogeneity by incorporating a mechanism that allows the analyst to bypass the assumption about the parameter distributions, which may not always be consistent across observations (compared to random parameter approaches; see Mannering et al., 2016). In addition, latent class models can account for unobserved heterogeneity through the assumption that observations come from classes that are based on common characteristics and that are distinct in nature (Mannering et al., 2016). The latent class order probit model will be developed for both rural and urban contexts using crash data pertaining to ROR crashes involving large trucks in Oregon between 2007 and 2014. This study contributes to the body of knowledge in the context of large truck safety by narrowing the gap in the literature regarding the influence of land use setting on the injury severity of ROR crashes. To the best of the authors’ knowledge, this study is among the first attempts to identify potential contributing factors to the injury severity of ROR crashes involving large trucks in urban and rural areas.

**4.2 Literature Review**

From a methodological perspective, a wide variety of approaches have been applied in the study of injury severity of roadway crashes involving large trucks (Al-Bdairi et al., 2018; Al-Bdairi and Hernandez, 2017; Anderson and Hernandez, 2017; Chen and Chen, 2011; Islam, 2015; Islam and Hernandez, 2013a; Islam et al., 2014; Khorashadi
et al., 2005; Lemp et al., 2011; Osman et al., 2016; Pahukula et al., 2015). Most of these studies have investigated the injury severity of truck drivers by utilizing random parameter models, with the exception of Khorashadi et al. (2005) and Lemp et al. (2011), who applied fixed parameter models. Even though past studies show that there is a growing interest in utilizing random parameter models in analyzing the injury severity of large truck drivers, still this approach of accounting for unobserved heterogeneity has some restrictions. For instance, in random parameter models, unobserved heterogeneity is accounted for through the underlying assumption that the estimated parameters vary across observations according to some predefined distributional forms. The issue with this approach is that the analyst prespecifies the distributions (this can be difficult depending on the analyst’s level of understanding) needed to estimate the potential random parameters (Washington et al., 2011). In contrast, latent class models (finite mixture) free the analyst from prespecifying the underlying distributional forms by identifying subgroups of data relying upon a specified number of mass points (Mannering et al., 2016; Mannering and Bhat, 2014). Lastly, recent advancements in econometric methods have begun utilizing methodologies that combine random parameters along with latent class models to better account for unobserved heterogeneity in crash data (Xiong and Mannering, 2013). For this study, we utilize a latent class approach on ROR crashes for two subpopulations of land use type, namely urban and rural areas.

Although not directly related, there are a number of studies that have disaggregated data into subpopulations (e.g., by age, gender, time of day, roadway classification, area type, light condition, etc.) to study the effect of those subpopulations on injury severities sustained by large truck drivers. For instance, Islam and Mannering (2006) disaggregated crash data into six models for three age groups and for both genders. Morgan and Mannering (2011) followed the same approach by separating driver age into two subpopulations and pavement conditions into three subpopulations (dry, wet, snow/icy). To study the impact of alcohol impairment on large truck crashes, Behnood et al. (2014) investigated the injury severity of alcohol-impaired drivers and those who
were sober at the time of the crash by splitting the data according to alcohol-impairment status, driver age, and gender. Pahukula et al. (2015) analyzed the injury severity of heavy vehicle crashes by separating crash data by time of day. Anderson and Hernandez (2017) analyzed the injury severity of heavy vehicle drivers based on roadway classifications. In addition, roadway lighting has been disaggregated to study its effect on the injury severity sustained by large truck drivers (Al-Bdairi et al., 2018; Anarkooli and Hosseinlou, 2016).

However, studies that attempt to quantify the impacts of land use on ROR crashes involving large trucks are sparse. For example, Khorashadi et al. (2005) conducted a study to highlight the differences in driver injury severities between urban and rural settings for crashes involving large trucks by using multinomial logit models. This study, however, ignores the effect of unobserved heterogeneity on injury severities. Moreover, Khorashadi et al. (2005) examined all types of truck-related crashes rather than emphasizing ROR crashes. In a different vein, Islam et al. (2014) investigated the effects of area type on injury severity, along with the number of vehicles involved in at-fault large truck crashes, by developing a mixed logit model. Still, the main focus of this study was on at-fault, large truck-related crashes.

In summary, past studies have characteristically utilized random parameter modeling frameworks to analyze injury severities, but these studies may be restrictive in that the analyst prespecified the random parameter distributions, in contrast to the latent class approach. Furthermore, the study of land use type separately (urban vs. rural) as a subpopulation and a contributing factor in a holistic model is sparse, especially from the perspective of ROR crashes involving large trucks. Studying this aspect of ROR injury severities can provide greater insight into the contributing factors of these types of crashes in specific urban and rural area contexts.
4.3 Data Description

The analyses in this study were conducted using eight years of police-reported crash data regarding ROR crashes involving large trucks that occurred in Oregon between 2007 and 2014. The crash data that is maintained by the Oregon Department of Transportation (ODOT) includes detailed information about the characteristics of the crashes, the drivers involved, the environmental conditions, and the roadway inventory. The crash data was filtered to include only maximum injury severity sustained by drivers, and, in total, 3,054 crashes were included in this study for this particular time period. Given that the effect of urban and rural settings on the injury severity of ROR crashes involving large trucks is of particular interest in the current study, the crash data was further split into two datasets: one for ROR crashes involving large trucks that occurred in rural areas, with 2,253 (74% of crashes) observations, and the other one pertaining to crashes that took place in urban areas, with 801 (26% of crashes) observations.

The injury severity levels in the crash data are categorized as: no injury, possible injury, incapacitating injury, non-incapacitating injury, and fatal injury. In this study, however, injury severity is collapsed into four main ordered categories: no injury, possible injury, moderate injury (non-incapacitating), and severe injury (incapacitating and fatal) so that the number of observations and the percentage of each level of injury severity is sufficient for the analyses. A frequency and percentage distribution of driver injury severity of ROR crashes involving large trucks in urban and rural areas is depicted in Table 4.1. As this table clearly shows, ROR crashes involving large trucks occurring in urban areas are less severe than those occurring in rural areas. For instance, 84.1% of ROR crashes in urban areas resulted in a no injury outcome, whereas 65.8% of ROR crashes in rural areas resulted in the same outcome. The explanatory variables that were found to be significant at a 95% confidence level are presented in Table 4.2. In total, 17 explanatory variables are included in the analyses. These variables are distributed as five significant and exclusive variables to the rural area model, five significant and exclusive variables to the urban counterpart model, and seven variables...
common for both models. Tables 4.3 and 4.4 illustrate the frequency and percentage distribution of selected explanatory variables for ROR crashes involving large trucks in both urban and rural models.

Table 4.1: Frequency and percentage distribution of driver injury in urban and rural

<table>
<thead>
<tr>
<th>Injury severity</th>
<th>Urban</th>
<th>Rural</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Observations</td>
<td>Percent</td>
</tr>
<tr>
<td>Severe injury (fatal &amp; incapacitating)</td>
<td>15</td>
<td>1.9%</td>
</tr>
<tr>
<td>Moderate (non-incapacitating)</td>
<td>45</td>
<td>5.6%</td>
</tr>
<tr>
<td>Possible injury</td>
<td>67</td>
<td>8.4%</td>
</tr>
<tr>
<td>No injury</td>
<td>674</td>
<td>84.1%</td>
</tr>
<tr>
<td>Total observations</td>
<td>801</td>
<td>-</td>
</tr>
</tbody>
</table>
Table 4.2: Description of selected explanatory variables used in the analyses of urban and rural models

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description of variables</th>
<th>Effect of variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>DITCH</td>
<td>Harmful event (1 for colliding with ditch, 0 otherwise)</td>
<td>✓</td>
</tr>
<tr>
<td>GRADE</td>
<td>Roadway characteristics (1 for vertical curve, 0 otherwise)</td>
<td>✓</td>
</tr>
<tr>
<td>POP25K</td>
<td>Population density (1 if between 10,001 and 25,000, 0 otherwise)</td>
<td>✓</td>
</tr>
<tr>
<td>DRY</td>
<td>Roadway surface condition (1 for dry, 0 otherwise)</td>
<td>✓</td>
</tr>
<tr>
<td>LOSTCTRL</td>
<td>Losing control of vehicle (1 for yes, 0 otherwise)</td>
<td>✓</td>
</tr>
<tr>
<td>DLIT</td>
<td>Lighting condition (1 if darkness with street lights, 0 otherwise)</td>
<td>✓</td>
</tr>
<tr>
<td>NOSPED</td>
<td>Exceeding the posted speed or driving too fast for conditions (1 for no, 0 otherwise)</td>
<td>✓</td>
</tr>
<tr>
<td>NOBELT</td>
<td>Driver safety seatbelt (1 if not used, 0 otherwise)</td>
<td>✓</td>
</tr>
<tr>
<td>FALL</td>
<td>Month of the year (1 if between September and December, 0 otherwise)</td>
<td>✓</td>
</tr>
<tr>
<td>FATIGUE</td>
<td>Driver was fatigued (1 for yes, 0 otherwise)</td>
<td>✓</td>
</tr>
<tr>
<td>CURVE</td>
<td>Roadway characteristics (1 for horizontal curve, 0 otherwise)</td>
<td>✓</td>
</tr>
<tr>
<td>RSDMEDN</td>
<td>Median type (1 for raised median, 0 otherwise)</td>
<td>✓</td>
</tr>
<tr>
<td>MPH55</td>
<td>Speed limit (1 if 55 mph, 0 otherwise)</td>
<td>✓</td>
</tr>
<tr>
<td>OVRTURN</td>
<td>Harmful event (1 for overturn, 0 otherwise)</td>
<td>✓</td>
</tr>
<tr>
<td>NODEPLOY</td>
<td>Airbag deployment (1 if did not deploy, 0 otherwise)</td>
<td>✓</td>
</tr>
<tr>
<td>YOUNG</td>
<td>Driver age (1 if between 20 and 45 years, 0 otherwise)</td>
<td>✓</td>
</tr>
<tr>
<td>FEMALE</td>
<td>Driver gender (1 if female, 0 otherwise)</td>
<td>✓</td>
</tr>
</tbody>
</table>
Table 4.3: Frequency and percentage distribution of explanatory variables in urban model

<table>
<thead>
<tr>
<th>Variable</th>
<th>Severe injury</th>
<th>Moderate injury</th>
<th>Possible injury</th>
<th>No injury</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>DRY</td>
<td>13 (2.4%)</td>
<td>35 (6.6%)</td>
<td>43 (8.1%)</td>
<td>440 (82.9%)</td>
<td>531</td>
</tr>
<tr>
<td>DLIT</td>
<td>3 (3.2%)</td>
<td>4 (4.3%)</td>
<td>12 (2.9%)</td>
<td>74 (79.6%)</td>
<td>93</td>
</tr>
<tr>
<td>DITCH</td>
<td>2 (3.3%)</td>
<td>8 (13.3%)</td>
<td>9 (15.0%)</td>
<td>41 (68.3%)</td>
<td>60</td>
</tr>
<tr>
<td>FATIGUE</td>
<td>0 (0.0%)</td>
<td>4 (20.0%)</td>
<td>5 (25.0%)</td>
<td>11 (55.0%)</td>
<td>20</td>
</tr>
<tr>
<td>NOSPED</td>
<td>8 (1.3%)</td>
<td>24 (4.0%)</td>
<td>38 (6.4%)</td>
<td>524 (88.2%)</td>
<td>594</td>
</tr>
<tr>
<td>CURVE</td>
<td>2 (1.8%)</td>
<td>10 (9.2%)</td>
<td>17 (15.6%)</td>
<td>80 (73.4%)</td>
<td>109</td>
</tr>
<tr>
<td>NOBELT</td>
<td>4 (12.9%)</td>
<td>7 (22.6%)</td>
<td>5 (16.1%)</td>
<td>15 (48.4%)</td>
<td>31</td>
</tr>
<tr>
<td>FALL</td>
<td>5 (2.3%)</td>
<td>13 (5.9%)</td>
<td>21 (9.5%)</td>
<td>183 (82.4%)</td>
<td>222</td>
</tr>
<tr>
<td>GRADE</td>
<td>2 (3.8%)</td>
<td>8 (15.4%)</td>
<td>10 (19.2%)</td>
<td>32 (61.5%)</td>
<td>52</td>
</tr>
<tr>
<td>RSDMEDN</td>
<td>5 (3.6%)</td>
<td>5 (3.6%)</td>
<td>14 (10.1%)</td>
<td>114 (82.6%)</td>
<td>138</td>
</tr>
<tr>
<td>POP25K</td>
<td>1 (0.7%)</td>
<td>4 (2.9%)</td>
<td>8 (5.8%)</td>
<td>124 (90.5%)</td>
<td>137</td>
</tr>
<tr>
<td>LOSTCTRL</td>
<td>5 (3.4%)</td>
<td>16 (11.0%)</td>
<td>24 (16.6%)</td>
<td>100 (69.0%)</td>
<td>145</td>
</tr>
</tbody>
</table>

Table 4.4: Frequency and percentage distribution of explanatory variables in rural model

<table>
<thead>
<tr>
<th>Variable</th>
<th>Severe injury</th>
<th>Moderate injury</th>
<th>Possible injury</th>
<th>No injury</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>MPH55</td>
<td>44 (3.2%)</td>
<td>252 (18.2%)</td>
<td>221 (16.0%)</td>
<td>867 (62.6%)</td>
<td>1384</td>
</tr>
<tr>
<td>NODEPLOY</td>
<td>11 (2.2%)</td>
<td>71 (14.3%)</td>
<td>71 (14.3%)</td>
<td>342 (69.1%)</td>
<td>495</td>
</tr>
<tr>
<td>OVTURN</td>
<td>10 (2.4%)</td>
<td>108 (26.1%)</td>
<td>85 (20.5%)</td>
<td>211 (51.0%)</td>
<td>414</td>
</tr>
<tr>
<td>FEMALE</td>
<td>5 (5.2%)</td>
<td>22 (22.7%)</td>
<td>19 (19.6%)</td>
<td>51 (52.6%)</td>
<td>97</td>
</tr>
<tr>
<td>YOUNG</td>
<td>25 (2.3%)</td>
<td>140 (13.0%)</td>
<td>174 (16.2%)</td>
<td>738 (68.5%)</td>
<td>1077</td>
</tr>
<tr>
<td>DRY</td>
<td>54 (4.6%)</td>
<td>259 (21.9%)</td>
<td>165 (14.0%)</td>
<td>702 (59.5%)</td>
<td>1180</td>
</tr>
<tr>
<td>DLIT</td>
<td>1 (1.8%)</td>
<td>4 (7.0%)</td>
<td>10 (17.5%)</td>
<td>42 (73.7%)</td>
<td>57</td>
</tr>
<tr>
<td>DITCH</td>
<td>36 (7.0%)</td>
<td>113 (22.1%)</td>
<td>88 (17.2%)</td>
<td>274 (53.6%)</td>
<td>511</td>
</tr>
<tr>
<td>FATIGUE</td>
<td>3 (2.5%)</td>
<td>41 (34.5%)</td>
<td>21 (17.6%)</td>
<td>54 (45.4%)</td>
<td>119</td>
</tr>
<tr>
<td>NOSPED</td>
<td>41 (3.7%)</td>
<td>192 (17.3%)</td>
<td>145 (13.1%)</td>
<td>733 (66.0%)</td>
<td>1111</td>
</tr>
<tr>
<td>CURVE</td>
<td>34 (4.4%)</td>
<td>170 (22.1%)</td>
<td>123 (16.0%)</td>
<td>442 (57.5%)</td>
<td>769</td>
</tr>
<tr>
<td>NOBELT</td>
<td>24 (17.5%)</td>
<td>52 (38.0%)</td>
<td>28 (20.4%)</td>
<td>33 (24.1%)</td>
<td>137</td>
</tr>
</tbody>
</table>
4.4 Methodology

To analyze and determine the potential factors contributing to the injury severity (ordered response) of particular crashes, ordered-response discrete choice models such as ordered probit/logit models are commonly utilized (Abdel-Aty, 2003; Al-Bdairi and Hernandez, 2017; Anarkooli and Hosseinalou, 2016; Haleem and Abdel-Aty, 2010; Osman et al., 2016; Quddus et al., 2002; Zhu and Srinivasan, 2011). Despite the abundance of studies that have examined driver injury severity in the transportation safety context, it is surprising that studies employing latent class ordered probit models are sparse. Alternatively, random parameter discrete choice models for ordered-response variables have been extensively used in previous studies (Al-Bdairi and Hernandez, 2017; Christoforou et al., 2010; Islam and Hernandez, 2013a; Naik et al., 2016). The latent class model is an alternative way to address the heterogeneity in injury severity analyses and is utilized in the current study (Mannering et al., 2016). The latent class approach accounts for possible unobserved heterogeneity present in the process of the data collection and the data itself, and it provides a means by which an analyst can bypass the assumptions about the parameter distributions, which may not always be consistent across observations (compared to random parameter approaches; see Mannering et al., 2016). In addition, the latent class modeling framework can account for unobserved heterogeneity through the assumption that observations come from classes that are based on common characteristics and that are distinct in nature (Mannering et al., 2016).

Regarding the latent class ordered probit model used in this analysis to analyze the injury severity of large truck drivers in ROR crashes, it is assumed that large truck drivers are distributed into C homogenous classes based on the characteristics of ROR crashes. It should be noted that an analyst does not know from the crash data from which class an observation is drawn. Moreover, each class has its own explanatory factors (Greene and Hensher, 2010).
Within each class, the contributing factors that affect driver injury severity involving ROR crashes are assumed to be fixed. As such, a traditional ordered probit model is used to estimate those factors within each class. Let \( c \) be the number of classes (\( c = 1, 2, \ldots, C \)), \( i \) represents the index for drivers (\( i = 1, 2, \ldots, N \)), and \( j \) is the driver injury severity out of \( J \) total injury severity outcomes (\( j = 1, 2, \ldots, J \)). To derive the traditional ordered probit model, an underlying continuous utility function \( y_i^* \) should be defined. This function can be used to determine the discrete injury severity outcomes of truck drivers involved in ROR crashes conditional on driver \( i \) belonging to class \( c \), as written in Eq. (4.1) (Washington et al., 2011).

\[
y_i^*[i \in c] = X_i \beta_c + \epsilon_{ic}, y_{ic} = j, \text{if } \mu_{i,j-1,c} < y_i^* < \mu_{i,j,c}
\]

(4.1)

where \( X_i \) is a vector of explanatory variables that contribute to driver injury severity; \( \beta_c \) is the associated vector of estimable parameters that belong to class \( c \); \( \epsilon_{ic} \) is an error term or a disturbance term, which is assumed to be independently randomly distributed; and \( \mu_{i,j,c} \) denotes the upper threshold associated with a particular class \( c \) that defines the injury severity outcome \( j \) for a driver \( i \) (Yasmin et al., 2014). Now, to determine the probability that large truck driver \( i \) sustains injury severity outcome \( j \) when involved in ROR crashes conditional on driver \( i \) belonging to class \( c \), Greene and Hensher (2010) and Yasmin et al. (2014) illustrate this as follows:

\[
P_i(j)|c = \Phi (\mu_{i,j,c} - X_i \beta_c) - \Phi (\mu_{i,j-1,c} - X_i \beta_c)
\]

(4.2)

where \( \Phi(.) \) denotes the standard normal cumulative distribution function for the error term. Since the analyst is unable to know which class a driver \( i \) belongs to, a vector representing observed crash factors \( \eta_i \) is utilized to identify that class. Greene and Hensher (2010) proposed using a multinomial logit structure to determine the
probability of assigning a driver $i$ to class $c$ while forcing the class probabilities to be between zero and one and to sum to one, as shown in Eq. (4.3).

$$P_{ic} = \frac{\exp(\alpha_c \eta_i)}{\sum_c \exp(\alpha_c \eta_i)}$$

(4.3)

where $\alpha_c$ is a vector of estimated parameters. Next, to determine the unconditional probability of driver $i$ sustaining injury severity $j$, Yasmin et al. (2014) used the formula illustrated in Eq. (4.4). Finally, Eq. (4.5) represents the log-likelihood function for the entire dataset (Yasmin et al., 2014).

$$P_i(j) = \sum_{c=1}^{C} (P_i(j)|c) \times (P_{ic})$$

(4.4)

$$L = \sum_{i=1}^{N} \log \left[ \sum_{c=1}^{C} (P_i(j)|c) \times (P_{ic}) \right]$$

(4.5)

Together, these equations provide a flexible methodology by which the injury severity of ROR crashes involving large trucks on urban and rural roadways can be studied. Lastly, marginal effects are computed for a better interpretation of the results and to determine the effect of each explanatory parameter on the injury severity outcome probabilities. Since indicator variables are created in this study, the marginal effects represent the numerical difference of the injury severity outcome probabilities, while the indicator variables change from zero to one (Washington et al., 2011).

$$M^{P_i(j)}_{X_{ij}} = P_i(j)(\text{given } X_{ij} = 1) - P_i(j)(\text{given } X_{ij} = 0)$$

(4.6)
4.4.1 Model separation tests

Despite the substantial differences between injury severities sustained by large truck drivers involved in ROR crashes in rural and urban areas, robust statistical methods need to be used to highlight those differences, as well as the commonalities between the two models. Such methods can validate if separate models should be developed for rural and urban ROR crashes over one aggregated model. To achieve this, two series of likelihood ratio tests are commonly used. The first is the log-likelihood ratio test that can be conducted to examine if the effects of the contributing factors that are identified to have a direct association with injury severities in two separated models (rural and urban) are similar to those in a holistic model that combines rural and urban crashes as a whole (Washington et al., 2011). In this test, it is hypothesized that the holistic model and the separate models are the same from the perspective of the contributing factors in the sense that the difference is not statistically significant. The first log-likelihood ratio test is illustrated in Eq. (4.7) (Washington et al., 2011).

\[ \chi^2 = -2[LL(\beta_H) - LL(\beta_U) - LL(\beta_R)] \]  

(4.7)

where \( LL(\beta_H) \) is the log-likelihood at convergence for the holistic model that combines ROR crashes occurring in both rural and urban areas; \( LL(\beta_U) \) is the log-likelihood at convergence for the urban model; and \( LL(\beta_R) \) represents the log-likelihood at convergence for the rural model. In this study, the obtained value of log-likelihood at convergence for the holistic model is \(-2479.445\), with 17 estimated parameters (degrees of freedom), while the log-likelihood at convergence for each developed separate model is provided in Table 4.6. The chi-square statistic obtained after applying Eq. (4.7) is distributed with seven degrees of freedom (the total number of estimated parameters in both the rural and urban model minus those in the holistic model). As shown in Table 4.6, the values of log-likelihood at convergence for rural and urban models are \(-1989.500\) and \(-391.102\), respectively. That is, the chi-square statistic \( \chi^2 \)
determined by Eq. (4.7) is equal to 197.686 with seven corresponding degrees of freedom. Accordingly, the null hypothesis that there is an insignificant statistical difference between the holistic model and the separate models as regards to the contribution factors must be rejected with well over 99% confidence, meaning that the models for ROR crashes involving large trucks in rural and urban areas must be developed and estimated separately.

The second log-likelihood ratio test is the parameter transferability test by which the stability of parameter estimates can be tested. This test is another approach that justifies using separate models in terms of ROR crashes in rural and urban areas in lieu of a holistic one that combines crashes in both areas. Washington et al. (2011) formulate the parameter transferability test as written in Eq. (4.8).

$$
\chi^2 = -2[LL(\beta_{ba}) - LL(\beta_a)]
$$

(4.8)

where $LL(\beta_a)$ is the log-likelihood at convergence for model $a$ (i.e., rural model) that is estimated based on $a$’s data (rural data) without any restriction, and $LL(\beta_{ba})$ is the log-likelihood at convergence for model $a$ (rural model) using the converged parameters from model $b$ (urban model). This test was also reversed in this study. Applying Eq. (4.8) can yield a chi-square statistic $\chi^2$ that follows a chi-square distribution with degrees of freedom equal to the number of estimated parameters in $LL(\beta_{ba})$. Therefore, restricting the estimated parameters in the urban model to be the rural estimated parameters, and vice versa by applying Eq. (4.8). By doing so, the values of the chi-square statistics for both cases are 695 and 181, respectively. These values with seven corresponding degrees of freedom indicate once again that with over 99% confidence, the developed separate models representing ROR crashes involving large trucks in rural and urban areas have different estimated parameters. Accordingly, developing two separate models for area type is justified.
4.5 Estimation results

As mentioned earlier, the formulation of a latent class model entails specifying the number of homogeneous classes or groups of observations with similar impacts on injury severity (Mannering et al., 2016). Yet, such classification is usually associated with challenges, from a computation perspective, when the number of classes or groups increases. Instead, the number of classes should be determined by the analyst. In this study, two classes were used in the analyses for latent class ordered probit models that were developed to estimate the impacts of contributing factors on injury severity of ROR crashes involving large trucks in urban and rural areas. Table 4.5 shows the estimated probabilities and shares of severity outcomes of each class for both models.

Table 4.5 demonstrates that the probability of large truck drivers involved in ROR crashes on urban roadways being assigned to class 1 is higher (58.0%) than the probability of their being assigned to class 2. However, this distribution regarding the probability of assigning drivers to classes is completely reversed for ROR crashes occurring on rural roadways in the sense that the likelihood of drivers being assigned to class 2 is higher (51.0%) than their being assigned to class 1. In terms of the distribution of probabilities of injury severities sustained by truck drivers on both urban and rural roadways belonging to a specific class, Table 4.5 shows that drivers belonging to class 1 are more likely to sustain no injuries when they are involved in ROR crashes on both urban and rural roadways, with 92.2% and 83.0%, respectively. Conversely, the injury severities of drivers assigned to class 2 are substantially high in both models. In the urban model, the probabilities of severe, moderate, and possible injuries of drivers belonging to class 2 are 10.8%, 22.5%, and 28.3%, respectively. The same observation is valid for the rural model, in which drivers belonging to class 2 are more likely to sustain severe, moderate, and possible injuries with 9.1%, 31.0%, and 26.8%, respectively.
Table 4.5: Estimated probabilities for each latent class in urban and rural models

<table>
<thead>
<tr>
<th>Components</th>
<th>Urban Latent Class 1</th>
<th>Urban Latent Class 2</th>
<th>Rural Latent Class 1</th>
<th>Rural Latent Class 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Crash population share</td>
<td>0.580</td>
<td>0.420</td>
<td>0.490</td>
<td>0.510</td>
</tr>
<tr>
<td>Injury severity outcomes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Severe injury</td>
<td>0.003</td>
<td>0.108</td>
<td>0.003</td>
<td>0.091</td>
</tr>
<tr>
<td>Moderate</td>
<td>0.026</td>
<td>0.225</td>
<td>0.087</td>
<td>0.310</td>
</tr>
<tr>
<td>Possible injury</td>
<td>0.048</td>
<td>0.283</td>
<td>0.079</td>
<td>0.268</td>
</tr>
<tr>
<td>No injury</td>
<td>0.922</td>
<td>0.383</td>
<td>0.830</td>
<td>0.331</td>
</tr>
</tbody>
</table>

Since the latent class ordered probit models with two distinct classes were developed in this study, each variable in the analyses has two estimated parameters, as clearly shown in Table 4.6. Further, Table 4.6 reveals that some parameters have different signs across the two classes (i.e., fatigued drivers in both models), others have similar signs across classes (i.e., dry surface in both models), and other parameters are significant only in a specific class. Such variation in the effects of the parameters on injury severity incurred by large truck drivers involved in ROR crashes across classes indicates that there is a significant heterogeneity between the two classes. Besides, the latent class ordered probit model used in this study is an extension of the traditional ordered probit model, in which the impact of an estimated parameter on increasing or decreasing the probability of extreme ordered discrete injury severity levels (in this study, severe injury and no injury) is clear, while the effect of that parameter on the probability of intermediate injury levels (moderate and possible injuries) is ambiguous (Washington et al., 2011). Consequently, the interpretation of the findings will be based mainly on the marginal effects.

This study sought to identify factors contributing to the injury severity of large truck drivers involved in ROR crashes occurring in urban and rural areas. Thus, the crash dataset was split into two distinct datasets, one representing ROR crashes occurring in
urban areas and the second one for those occurring in rural areas. Consequently, two separate latent class ordered probit models were developed for ROR crashes in urban and rural areas, respectively. Table 4.6 presents the estimation results for both models. The marginal effects that were used to assess the effect of the estimated parameters in the urban and rural models are shown in Tables 4.7 and 4.8, respectively. Clearly, Table 4.6 shows that, in each model, 12 factors were found to be statistically significant in impacting the injury severity incurred by drivers. Also, seven factors were found to be significant in both models. To ease the interpretations of the study findings, the discussion of results will be presented in three subsequent sections. The first section will discuss the mutual factors in both models, while the second and third sections will highlight the exclusive factors that were found to affect injury severity in the urban and rural models, respectively.
Table 4.6: Estimation results of latent class ordered probit for urban and rural models

<table>
<thead>
<tr>
<th>Variable</th>
<th>Urban</th>
<th>Rural</th>
<th>Urban</th>
<th>Rural</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Parameter</td>
<td>Parameter</td>
<td>Parameter</td>
<td>Parameter</td>
</tr>
<tr>
<td></td>
<td>estimate</td>
<td>estimate</td>
<td>estimate</td>
<td>estimate</td>
</tr>
<tr>
<td></td>
<td>t-stat</td>
<td>t-stat</td>
<td>t-stat</td>
<td>t-stat</td>
</tr>
<tr>
<td>Constant</td>
<td>5.347</td>
<td>5.61</td>
<td>3.112</td>
<td>6.51</td>
</tr>
<tr>
<td>DITCH</td>
<td>0.355</td>
<td>0.57</td>
<td>-0.947</td>
<td>-1.94</td>
</tr>
<tr>
<td>GRADE</td>
<td>-1.597</td>
<td>-3.49</td>
<td>-0.222</td>
<td>-0.51</td>
</tr>
<tr>
<td>POP25K</td>
<td>-1.282</td>
<td>-2.38</td>
<td>8.564</td>
<td>0.00</td>
</tr>
<tr>
<td>DRY</td>
<td>-1.270</td>
<td>-2.54</td>
<td>-0.738</td>
<td>-2.67</td>
</tr>
<tr>
<td>LOSTCTRL</td>
<td>-1.413</td>
<td>-2.87</td>
<td>-0.361</td>
<td>-1.15</td>
</tr>
<tr>
<td>DLIT</td>
<td>-1.546</td>
<td>-2.93</td>
<td>-0.087</td>
<td>-0.21</td>
</tr>
<tr>
<td>NOSPED</td>
<td>2.417</td>
<td>3.76</td>
<td>-0.133</td>
<td>-0.33</td>
</tr>
<tr>
<td>NOBELT</td>
<td>-1.785</td>
<td>-2.33</td>
<td>-2.015</td>
<td>-2.93</td>
</tr>
<tr>
<td>FALL</td>
<td>-0.870</td>
<td>-2.35</td>
<td>0.266</td>
<td>0.93</td>
</tr>
<tr>
<td>FATIGUE</td>
<td>-4.069</td>
<td>-4.13</td>
<td>1.802</td>
<td>0.44</td>
</tr>
<tr>
<td>CURVE</td>
<td>1.125</td>
<td>1.70</td>
<td>-1.165</td>
<td>-2.57</td>
</tr>
<tr>
<td>RSDMEDN</td>
<td>1.250</td>
<td>2.07</td>
<td>-0.586</td>
<td>-1.89</td>
</tr>
<tr>
<td>MPH55</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>OVTURN</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>NODEPLOY</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>YOUNG</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>FEMALE</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Threshold 1</td>
<td>1.709</td>
<td>3.24</td>
<td>0.821</td>
<td>4.98</td>
</tr>
<tr>
<td>Threshold 2</td>
<td>2.986</td>
<td>4.31</td>
<td>1.431</td>
<td>6.72</td>
</tr>
<tr>
<td>Class Probability (t-stat)</td>
<td>0.58 (6.73)</td>
<td>0.42 (4.96)</td>
<td>0.49 (6.28)</td>
<td>0.51 (6.59)</td>
</tr>
<tr>
<td>Number of Observations</td>
<td>801</td>
<td>2253</td>
<td>801</td>
<td>2253</td>
</tr>
<tr>
<td>Log likelihood at convergence</td>
<td>-391.102</td>
<td>-1989.500</td>
<td>-471.823</td>
<td>-2176.024</td>
</tr>
<tr>
<td>Log likelihood at zero</td>
<td>-471.823</td>
<td>-2176.024</td>
<td>0.171</td>
<td>0.086</td>
</tr>
</tbody>
</table>
Table 4.7: Estimated marginal effects for latent class ordered probit of urban model

<table>
<thead>
<tr>
<th>Variable</th>
<th>Severe injury</th>
<th>Moderate injury</th>
<th>Possible injury</th>
<th>No injury</th>
</tr>
</thead>
<tbody>
<tr>
<td>DITCH</td>
<td>0.0116</td>
<td>0.0351</td>
<td>0.0427</td>
<td>-0.0894</td>
</tr>
<tr>
<td>GRADE</td>
<td>0.0308</td>
<td>0.0769</td>
<td>0.0810</td>
<td>-0.1887</td>
</tr>
<tr>
<td>POP25K</td>
<td>-0.0055</td>
<td>-0.0212</td>
<td>-0.0319</td>
<td>0.0586</td>
</tr>
<tr>
<td>DRY</td>
<td>0.0107</td>
<td>0.0395</td>
<td>0.0578</td>
<td>-0.1080</td>
</tr>
<tr>
<td>LOSTCTR</td>
<td>0.0111</td>
<td>0.0347</td>
<td>0.0435</td>
<td>-0.0893</td>
</tr>
<tr>
<td>DLIT</td>
<td>0.0117</td>
<td>0.0355</td>
<td>0.0436</td>
<td>-0.0908</td>
</tr>
<tr>
<td>NOSPED</td>
<td>-0.0165</td>
<td>-0.0500</td>
<td>-0.0615</td>
<td>0.1280</td>
</tr>
<tr>
<td>NOBELT</td>
<td>0.0894</td>
<td>0.1560</td>
<td>0.1244</td>
<td>-0.3698</td>
</tr>
<tr>
<td>FALL</td>
<td>0.0021</td>
<td>0.0074</td>
<td>0.0101</td>
<td>-0.0196</td>
</tr>
<tr>
<td>FATIGUE</td>
<td>0.0402</td>
<td>0.0923</td>
<td>0.0907</td>
<td>-0.2232</td>
</tr>
<tr>
<td>CURVE</td>
<td>0.0079</td>
<td>0.0254</td>
<td>0.0324</td>
<td>-0.0657</td>
</tr>
<tr>
<td>RSDMEDN</td>
<td>-0.0008</td>
<td>-0.0028</td>
<td>-0.0039</td>
<td>0.0074</td>
</tr>
</tbody>
</table>

Table 4.8: Estimated marginal effects for latent class ordered probit of rural model

<table>
<thead>
<tr>
<th>Variable</th>
<th>Severe injury</th>
<th>Moderate injury</th>
<th>Possible injury</th>
<th>No injury</th>
</tr>
</thead>
<tbody>
<tr>
<td>DITCH</td>
<td>0.0271</td>
<td>0.0924</td>
<td>0.0383</td>
<td>-0.1578</td>
</tr>
<tr>
<td>DRY</td>
<td>0.0166</td>
<td>0.0677</td>
<td>0.0341</td>
<td>-0.1184</td>
</tr>
<tr>
<td>DLIT</td>
<td>-0.0082</td>
<td>-0.0380</td>
<td>-0.0216</td>
<td>0.0679</td>
</tr>
<tr>
<td>NOSPED</td>
<td>-0.0058</td>
<td>-0.0236</td>
<td>-0.0118</td>
<td>0.0412</td>
</tr>
<tr>
<td>NOBELT</td>
<td>0.1129</td>
<td>0.2181</td>
<td>0.0414</td>
<td>-0.3724</td>
</tr>
<tr>
<td>CURVE</td>
<td>0.0133</td>
<td>0.0513</td>
<td>0.0242</td>
<td>-0.0888</td>
</tr>
<tr>
<td>YOUNG</td>
<td>-0.0104</td>
<td>-0.0429</td>
<td>-0.0216</td>
<td>0.0749</td>
</tr>
<tr>
<td>FEMALE</td>
<td>0.0253</td>
<td>0.0811</td>
<td>0.0309</td>
<td>-0.1374</td>
</tr>
<tr>
<td>FATIGUE</td>
<td>0.0235</td>
<td>0.0765</td>
<td>0.0298</td>
<td>-0.1298</td>
</tr>
<tr>
<td>MPH55</td>
<td>0.0063</td>
<td>0.0261</td>
<td>0.0134</td>
<td>-0.0457</td>
</tr>
<tr>
<td>OVRTURN</td>
<td>0.0219</td>
<td>0.0759</td>
<td>0.0319</td>
<td>-0.1296</td>
</tr>
<tr>
<td>NODEPLOY</td>
<td>-0.0013</td>
<td>-0.0054</td>
<td>-0.0027</td>
<td>0.0094</td>
</tr>
</tbody>
</table>
4.5.1 Factors contributing to ROR crashes on both models

As mentioned previously, seven variables were found to be statistically significant in both models, meaning that these variables have a substantial impact on the injury severity sustained by truck drivers involved in ROR crashes, regardless of the area type, and despite the fact that there are remarkable differences between the urban and rural areas from the perspective of geometric design, traffic volume, speed limit, and driver behavior. These variables are driver fatigue, dry roadway surface condition, dark with street lights, colliding with a ditch, seatbelt not used, the presence of horizontal curves, and neither exceeding the posted speed nor driving too fast for the conditions.

With regard to the influence of fatigue on injury severity sustained by large truck drivers, this variable was found to be statistically significant in both the urban and rural models. Further, this variable has different signs across the two classes in the sense it is negative and significant in class 1, but positive and insignificant in class 2 for both models, meaning that this variable has a heterogeneous impact on injury severity. Tables 4.7 and 4.8 illustrate that ROR crashes involving fatigued large truck drivers are less likely to cause no injuries, with −0.2232 and −0.1298 for the urban and rural models, respectively. Instead, the factor of driver fatigue in these crashes would increase the probability of incurring higher injury severity levels, specifically moderate injuries, by 0.0923 and 0.0765 for the urban and rural models, respectively. This finding underscores the impact of fatigue of large truck drivers regardless of the area type because, in both models, the moderate injury severity will increase by slightly the same probability. In general, large truck drivers are characterized by some unique factors compared to passenger vehicle drivers, such as irregular schedules, long working hours, night driving, and economic pressures. All these factors increase the possibility of truck driver fatigue. Previous studies that primarily focused on large truck crashes also found that crashes involving fatigued drivers were associated with a higher level of severity. For example, Chen and Chen (2011) examined the differences between single and multiple vehicle crashes involving large trucks. In their study, fatigue was found to increase the probability of injury in multiple truck crashes.
compared to single truck crashes. Islam et al. (2014) conducted a study to capture the differences and commonalities in injury severity of large truck drivers in urban and rural areas. They found that the impact of driver fatigue on injury severity resulting from crashes occurring in rural areas was higher, while this impact was insignificant for crashes occurring in the counterpart urban areas. However, the studies conducted by Chen and Chen (2011) and Islam et al. (2014) did not examine the effect of fatigue of large truck drivers involved in ROR crashes on injury severity in urban and rural areas, as this study does. In relation to past studies that emphasized ROR crashes involving large trucks, the findings of this study support the findings of Al-Bdairi and Hernandez (2017) and Al-Bdairi et al. (2018), which indicate that non-incapacitating and possible injuries were more likely to be sustained in ROR crashes involving fatigued drivers. However, the findings of this study contradict the findings of the Peng and Boyle (2012) study, which found that fatigue leads to fatal ROR crashes.

Roadway surface conditions, specifically the variable representing dry surfaces, was found to affect injury severity in both models, with negative signs in both classes, as shown in Table 4.6. This finding suggests that this variable has homogeneous effects across classes. Tables 4.7 and 4.8 demonstrate that ROR crashes involving large trucks occurring on dry roadway surfaces, whether in rural or urban areas, are less likely to result in no injuries, by 0.1080 and 0.1184 for urban and rural areas, respectively. A possible explanation could be attributed to driver behavior in the sense that drivers may underestimate the risk of injury severity resulting from ROR crashes on dry roadway surfaces. As such, drivers tend to increase their driving speed on dry roadway surfaces. A similar finding was reported in previous studies, which was that crashes involving large trucks occurring on dry roadway surfaces were less likely to result in no injuries, while increasing fatalities and incapacitating injuries (Al-Bdairi et al., 2018; Al-Bdairi and Hernandez, 2017; Anderson and Hernandez, 2017; Islam and Hernandez, 2013b; Peng and Boyle, 2012).

In terms of the lighting conditions, the variable of dark with lighted streets was found to have different effects on injury severity based on the area type. Table 4.7
confirms that ROR crashes involving large trucks occurring on urban roadways in the dark, but with street light, would decrease the probability of no injuries by 0.0908. Conversely, ROR crashes occurring in the dark, but with street light, on the rural roadways would have a 0.0679 higher probability of resulting in no injuries, as shown in Table 4.8. This finding could be used to mobilize transportation agencies to deploy street lighting along hotspot locations on rural roadways such as horizontal curves. This finding is in line with previous studies that found that large truck crashes occurring in the dark, but with lighted streets, had a lower probability of resulting in no injuries, while increasing the likelihood of severe injuries (Islam and Hernandez, 2013b; Lemp et al., 2011; Zhu and Srinivasan, 2011). Further, Islam et al. (2014) examined crashes involving single and multiple large trucks in urban and rural areas. They found that crashes involving multiple trucks in urban areas in the dark, with no street light, would increase the probability of severe injury by 2.5%. Therefore, the finding of the current study regarding the effect on injury severity of driving in the dark, with street light, contradicts the study by Islam et al. (2014). However, the researchers in that study did not consider injury severity incurred by large truck drivers involved in ROR crashes as the variable of interest (dependent variable) in their analyses.

It is well recognized that the main goal of safety practitioners and transportation planners is to keep vehicles in their traveling lanes. In spite of this, however, ROR crashes constitute more than half the crashes in the U.S. To reduce the severity of injuries resulting from such crashes, a proactive strategy should be to relocate and/or redesign roadside objects, such as ditches. In the current study, the variable of ROR crashes involving large trucks colliding with ditches was found to affect injury severity in both the urban and rural models. Examining further, this variable has a positive sign and is insignificant in class 1, whereas it has a negative sign and is significant in class 2 for the urban model, suggesting that this variable has heterogeneous effects. Table 4.7 reveals that colliding with ditches in urban areas reduces the probability of a no injury outcome by 0.0894, while increasing the likelihood of resulting in a possible injury outcome by 0.0427. As regards the rural model, the variable of ROR crashes
involving large trucks colliding with ditches was found to have negative signs in both classes, indicating a homogeneous influence. Table 4.8 shows that a no injury outcome is 0.1578 less likely to be incurred when trucks collide with ditches in rural areas, while increasing the probability of a moderate injury level outcome by 0.0924. Careful examination of the effect of the variable of large trucks colliding with ditches in urban and rural models discloses that the impact of this variable in rural areas is higher than in urban areas and that injuries tend to be more severe as well. Qin et al. (2013) also found that crashes involving large trucks colliding with ditches had a higher probability of resulting in non-incapacitating and minor injuries. Such findings could aid in improving the design of ditches, back slopes, and flattening slopes, and in relocating hazardous objects to minimize crash severity.

Not surprisingly, not using seatbelts can dramatically increase the odds of large truck drivers being involved in fatal crashes. Therefore, transportation agencies should stringently enforce existing seatbelt laws and conduct public education campaigns to save drivers’ and other road users’ lives. In the current study, the variable of not using seatbelts was found to have negative signs and was statistically significant in the two classes for both models. Intuitively, this means that this variable has homogeneous influences on the injury severity of ROR crashes occurring in urban and rural areas in the sense that drivers who do not wear seatbelts would be at a high risk of sustaining severe injuries. In the urban model, not wearing seatbelts decreased the probability of no injuries for large truck drivers involved in ROR crashes by 0.3698, while increasing the likelihood of moderate injuries by 0.1560, as shown in Table 4.7. Table 4.8 reveals that large truck drivers who did not use seatbelts had a 0.3724 lower probability of incurring no injuries when they were involved in ROR crashes in rural areas, whereas the moderate injury outcome would increase by 0.2181. Obviously, a careful examination of the impacts of not wearing seatbelts on the safety of large truck drivers in urban and rural areas implies that the influence of this variable on ROR crashes involving large trucks in rural areas is higher than in urban areas. This variation could be attributed to the high speed that characterizes rural roadways. This finding agrees
with previous studies that found unbelted drivers are at a high risk of incurring severe and fatal injuries (Chen and Chen, 2011; Russo et al., 2014; Schneider et al., 2009). Specifically, Chen and Chen (2011) concluded that unbelted large truck drivers had a 47.1% and 37.1% higher probability of being in fatal crashes involving single and multiple trucks, respectively. Hence, the trucking industry could benefit from this finding by imposing strict enforcement of safety seatbelt usage on their drivers.

Roadway alignment and horizontal curves play a vital role in the occurrence of ROR crashes involving large trucks because the operational characteristics of trucks, such as weight and length, can create centrifugal forces that push trucks away from the curve, which in turn increases the odds of trucks being overturned. In this study, it was observed that the variable of the horizontal curve was statistically significant in both models. Moreover, it was found that ROR crashes involving large trucks occurring on horizontal curves in urban areas have heterogeneous impacts on injury severity because this variable has a positive sign in class 1 and a negative sign in class 2. Further explanation regarding the effect of this variable in the urban model could be achieved through the marginal effects shown in Table 4.7. Marginal effects disclose that large truck drivers involved in ROR crashes on horizontal curves in urban areas have a 0.0657 lower probability of being injured, while the likelihood of sustaining possible injuries increases by 0.0324. The variable of ROR crashes occurring on horizontal curves in rural areas has homogeneous effects across the two classes because it has a negative sign in both classes. Table 4.8 reveals that large truck drivers involved in ROR crashes on horizontal curves in rural areas are 0.0888 less likely to incur no injuries, while increasing their odds by 0.0513 of sustaining moderate injuries. Evidently, ROR crashes involving large trucks on horizontal curves in rural areas tend to be more severe than those taking place in the counterpart areas. A possible explanation could be attributed to the higher posted speed on rural roadways as opposed to urban roadways. Moreover, other behavioral factors could aggravate ROR crashes in rural areas such as unbelted drivers and fatigue. This finding can help guide safety practitioners to identify the hotspot locations, particularly on rural roadways, that experience a high number of
ROR crashes related to horizontal curves. Accordingly, appropriate safety countermeasures could be taken, such as better curve delineation and installing curve warning signs to alert drivers to upcoming curves. This finding is consistent with previous studies that found that large truck crashes occurring on horizontal curves were associated with non-incapacitating injuries (Al-Bdairi et al., 2018; Al-Bdairi and Hernandez, 2017; Lemp et al., 2011). However, other studies concluded that such crashes increased the probability of incapacitating and fatal injuries (Islam and Hernandez, 2013a; Naik et al., 2016).

Lastly, the factor of large truck drivers who abide by the speed limit was found to drastically increase the probability of sustaining no injuries when they are involved in ROR crashes in both urban and rural areas. The effect of this variable is more pronounced in the urban model, in which drivers abiding by the speed limit would be 0.1280 more likely to incur no injuries (see Table 4.7). In the rural model, this variable has a similar effect but to a lesser extent since a no injury outcome would be increased by 0.0412. Al-Bdairi and Hernandez (2017) have reported a similar finding.

4.5.2 Factors contributing to ROR crashes on urban roadways

Five estimated parameters were found to be statistically significant in affecting the injury severity of ROR crashes involving large trucks in the urban model. However, only three of these parameters will be discussed due to their higher marginal effects, namely, crashes occurring on vertical curves, urban areas with a population density of between 10,001 and 25,000, and losing control of a vehicle. Regarding ROR crashes on vertical curves, this variable was found to be significant in class 1 and insignificant in class 2, with a negative sign in the two classes, meaning that this variable has a heterogeneous influence. This variable has also been found to be heterogeneous in a study conducted by Al-Bdairi et al. (2018), in which ROR crashes involving large trucks occurring in dark conditions on vertical curves were found to be random in their injury outcomes. In this sense, the minor injury severity sustained by large truck drivers
involved in such crashes would increase by approximately 42.3%, while being involved in crashes resulting in minor injuries would be 57.7% less likely. Also, in this study, the variable of ROR crashes occurring on vertical curves has a 0.1887 lower likelihood of resulting in no injuries, while increasing the probability of possible injuries by 0.0810. A possible explanation could be the limited visibility and sight distance on vertical curves compared to straight roadways, which in turn increase the odds of ROR crashes involving large trucks of being less likely to result in no injuries. This finding is consistent with previous studies that found that large truck crashes on vertical curves would increase the likelihood of minor injuries (Al-Bdairi et al., 2018; Anderson and Hernandez, 2017).

Regarding the effect of population density on ROR crashes involving large trucks, the variable of urban areas with a population density of between 10,001 and 25,000 was found to be significant, with a negative sign in class 1, but it is insignificant with a positive sign in class 2, indicating that this variable has heterogeneous impacts. Table 4.7 shows that no injury outcomes would have a 0.0586 higher probability of being sustained by large truck drivers involved in ROR crashes occurring in urban areas with a population density of between 10,001 and 25,000. This finding indicates that large truck drivers tend to drive cautiously in urban areas with such a population density to avoid being involved in a crash. Next, the variable of losing control of a vehicle was found to be significant and has negative signs in both classes, meaning that this variable has fixed effects. Further, ROR crashes occurring due to losing control of vehicles were 0.0893 less likely to result in no injuries, and, in fact, this variable would increase the probability of possible injuries by 0.0435, as shown in Table 4.7. Al-Bdairi and Hernandez (2017) also found that ROR crashes involving large trucks that occur due to losing control of vehicles would be 95.5% less likely to cause no injuries. However, Al-Bdairi and Hernandez (2017) found that the variable of losing control of a vehicle was random, whereas this is not the case in this study.
4.5.3 Factors contributing to ROR crashes on rural roadways

In the rural model, the estimation results shown in Table 4.6 illustrate that five estimated parameters were exclusively significant in the rural model. However, only the parameters with higher impacts (higher marginal effects) will be presented here. Those parameters are young drivers between 20 and 45 years, female drivers, and overturning crashes. In terms of driver characteristics, two factors play a key role in ROR crashes involving large trucks in rural areas: driver age and female drivers. The indicator of young drivers has positive signs in the two classes and significantly impacts injury severity. Table 4.8 demonstrates that young drivers have a 0.0749 higher probability of sustaining no injuries when they are involved in ROR crashes in rural areas. This finding could be attributed to the physiological capabilities that characterize this age group, which make them more resilient in ROR crashes involving large trucks so that they are less likely to sustain severe injuries. Similarly, Anderson and Hernandez (2017) and Pahukula et al. (2015) found that large truck drivers between 35 and 45 years were more likely to sustain no injuries.

In addition to driver age, driver gender (female) was also found to be more vulnerable to severe injuries. Table 4.8 shows that female drivers had a 0.1374 lower probability of sustaining no injuries, while their likelihood of sustaining moderate injuries increased by 0.0811. The difference between males and females in incurring injuries may explain why female drivers are at a high risk of being involved in severe crashes as opposed to male drivers. This finding has been confirmed by Chen and Chen (2011) and Islam et al. (2014). With respect to overturning crashes, this event in rural areas is 0.1296 less likely to result in injuries, while it has a 0.0759 higher likelihood of causing moderate injuries, as shown in Table 4.8. This variable was found to be statistically significant in the two classes, with a negative sign in class 1 and a positive one in class 2, indicating heterogeneous influences. This finding is consistent with the Chen and Chen (2011) and Al-Bdairi et al. (2018) studies that found homogeneously that overturning crashes were associated with possible and non-incapacitating injuries in large truck drivers.
4.6 Summary and Conclusions

The present study analyzes the injury severity sustained by large truck drivers involved in ROR crashes at a disaggregate level for crashes occurring in urban and rural areas in the state of Oregon. Eight years of crash data pertaining to ROR crashes involving large trucks between 2007 and 2014 were used. In this study, injury severity was grouped into four main ordered categories: no injury, possible injury, moderate injury (non-incapacitating), and severe injury (incapacitating and fatal). Recognizing the ordinal nature of injury severity and the heterogeneity in crash data, latent class ordered probit models were developed to investigate the impact of land use setting on the injury severity of large truck drivers, while accounting for unobserved heterogeneity in crash data. Log-likelihood ratio tests were conducted to validate using separate models for urban and rural areas rather than a holistic model that combines them. The results of these tests indicate that the injury severity of ROR crashes occurring in urban and rural areas is quite different and therefore needs to be modeled separately.

The estimation results reveal that the developed latent class ordered probit models (for urban and rural areas) are substantially distinct in terms of the contributing factors. However, seven parameters were found to be significant in both the urban and rural models, but with remarkable differences in terms of their impacts on the injury severity incurred by large truck drivers involved in ROR crashes. For example, ROR crashes involving large trucks occurring on urban roadways at dark, but with street light, would decrease the probability of no injuries by 0.0908, whereas such crashes in the counterpart area would have a 0.0679 higher probability of resulting in no injuries. This variation underscores the substantial role of street lighting on rural roadways. As such, a safety countermeasure that could be recommended is the deployment of street lights along rural roadways in an attempt to reduce or avoid severe injuries resulting from
ROR crashes involving large trucks. Further, some variables were found to be exclusively significant in one model (urban or rural), but not both.

In each model, five estimated parameters were found to be uniquely significant. In the urban model, for example, ROR crashes occurring on vertical curves were found to decrease the no injury outcome by 0.1887 while increasing the probability of possible injuries by 0.0810. The limited visibility and sight distance on vertical curves could be underlying factors in this finding. Such a finding can motivate transportation agencies in Oregon to make further efforts to improve the vertical curves in urban areas in terms of visibility, sight distance, traffic control management, and posted speed limits. In the rural model, driver gender was found to play a vital role in injury severity outcomes. For instance, female drivers have a 0.1374 lower probability of sustaining no injuries, while their likelihood of sustaining moderate injuries increases by 0.0811. This could be related to the behavioral and physiological differences between male and female drivers such as driving experience, driving characteristics (i.e., aggression and risk perception), and the ability of the body to withstand impact. These differences between males and females in incurring injuries may explain why female drivers are at a high risk of being involved in severe crashes as opposed to male drivers.

To the best of the authors’ knowledge, utilizing separate latent class ordered probit models to analyze the injury severity of large truck drivers involved in ROR crashes in urban and rural areas is the first attempt to extend the literature of analyzing the injury severity of large truck crashes. It also fills the gap in the literature in terms of examining the injury severity of ROR crashes involving large trucks by land use setting (urban and rural). The trucking industry, transportation agencies, and safety practitioners could benefit from the findings of the current study to prevent or alleviate the injury severity of ROR crashes involving large trucks by developing appropriate and cost-effective countermeasures. However, despite the fact that the findings of this study can provide better understanding in terms of the factors contributing to the injury severity of large truck drivers involved in ROR crashes in each area type, some limitations should be pointed out. Some important factors that may contribute to ROR crashes are
missing in the Oregon crash data, such as pavement conditions, shoulder type, shoulder width, and roadside characteristics. Therefore, future studies should consider these limitations and aim to collect even more comprehensive crash data. Further, alternative advanced methods that address heterogeneity in the crash data, such as latent class models with random parameters within classes, could be used in future research.
References


Oregon Department of Transportation (ODOT), 2017. 2015 Oregon Traffic Crash Summary. Salem, Oregon.


Chapter 5 : Conclusions and Recommendations

5.1 Summary and Conclusions

The main objective of this dissertation was to identify contributing factors that significantly impact injury severity sustained by drivers of large trucks involved in ROR crashes. To do so, Oregon crash data for large truck crashes that took place between 2007 and 2013 were obtained ODOT. Given that the data are police crash reports, such data is characterized by some inherent limitations despite the fact that such crash reports can be useful in providing information that can be used in assessing a crash occurrence. Among the crash data limitations of note is the limited information that can be obtained by police officers at a crash scene from drivers involved in a crash. This is because those drivers either fail to remember the pre-crash details (e.g., falling asleep, being fatigued, being distracted) due to the crash shock (memory failure) or they may refrain to report crashes in which they were at fault such as being under the influence of alcohol or drugs. Another issue with police crash reports is the underreporting of low injury severity and/or property damage only (PDO) crashes. This underreporting could lead to substantial bias in injury severity prediction since the sample of crash data used in the analysis will not be a random one. This also causes an overrepresentation of crashes with higher injury severity levels. The aforementioned limitations lend themselves to the issue of unobserved heterogeneity in the sense that not all factors that might contribute to a particular crash would be seen (available) to the analyst.

Given these limitations, the methodological approaches that can account for such drawbacks were used in this research. Consequently, the main focus of the current study was to determine the factors contributing to ROR crashes involving large trucks while addressing the unobserved heterogeneity issue through employing advanced econometric methods, namely the random parameters ordered probit model, mixed logit model, and latent class ordered probit model. The current dissertation includes three standalone manuscripts, published in peer-reviewed journals.
In the first manuscript, random parameters ordered probit model was developed to achieve the first objective in the dissertation that aims to quantitatively determine the impacts of environmental conditions, roadway characteristics, driver characteristics, and vehicle factors on ROR crash injury severity sustained by truck drivers.

Manuscript two highlighted the differences and commonalities between lighted and dark conditions in regard to its influence on injury severity of large truck drivers involved in ROR crashes. To achieve this objective coupled with addressing unobserved heterogeneity, mixed logit models were used. Additionally, log-likelihood ratio tests (as model separation tests) were conducted to statistically validate if lighting conditions should be modeled separately or holistically. The model separation tests indicate with 99.99% confidence that lighting conditions should be modeled separately. Further evidence for considering lighting conditions separately in analyses rather than holistically was the large number of contributing factors that were found to exclusively affect a specific lighting condition.

Examination of the variation between rural and urban environments in regard to injury severity of large truck drivers involved in ROR crashes was presented in the third manuscript. Quantifying these variations would be useful in developing the appropriate safety countermeasures that could alleviate the number and severity of ROR crashes in rural/urban settings. With that in mind, the objective of the third manuscript was to pinpoint factors contributing to injury severity of large truck drivers involved in ROR crashes occurring in rural and urban areas. By doing so, a better understanding of disparities and commonalities in rural/urban injury severity was identified. An alternative way to random parameters method, that commonly used in accounting for unobserved heterogeneity, was applied, namely latent class ordered probit model for each land use setting. Once again, model separation tests were also performed for statistical validation of using a separate model for each land use setting in lieu of a holistic one that represents area type through indicator variables. Obtaining 99.99% confidence level in model separation tests makes one more confident that such separation should be performed to gain more insights.
5.2 Key findings

Based on the estimation results concerning injury severity of large truck drivers involved in ROR crashes obtained through utilizing random parameters and latent class models that commonly used in accounting for unobserved heterogeneity, findings of this dissertation are summarized.

1. ROR crashes occurring on roadways with raised medians were more likely to cause no injuries for drivers, particularly at dark conditions and urban roadways. Thurgood (2010) defined a raised median as a physical barrier installed in the center of a multilane roadway in which separation of opposite traffic is the primary objective. As such, head-on-collisions are eliminated via implementing raised median. The finding of this dissertation showed that raised medians highly increased the likelihood of no injuries of ROR crashes involving large trucks on dark conditions and urban roadways. A similar finding was found in a previous study conducted by Schultz et al. (2011) that found installing raised medians on Utah roadways reduced severe injury crashes by 36%. Yet, this finding is in conflict with Thurgood (2010), in which raised medians were recommended as an appropriate countermeasure on high volume and high-speed roadways while the finding of this dissertation demonstrates that raised median is a highly effective countermeasure on urban roadways. A possible explanation could be attributed to the role that raised medians can play in mitigating some types of collisions such as head-on and median crossover crashes through controlling access points on particular locations. This finding can provide conclusive evidence for transportation agencies and safety practitioners that installing raised medians on urban multilane roadways is an effective practice to reduce ROR crashes, and if such crashes occur, the severity outcomes would be minimal.
2. It is well recognized that driver fatigue can be associated with some implications, including a reduction in alertness, vigilance, performance, and motivation. Impaired judgment and feelings of drowsiness are other notorious consequences of driver fatigue (Knipling, 2015). The finding of this dissertation shows that fatigued drivers of large truck were more likely to incur non-incapacitating and possible injuries when they are involved in ROR crashes under all situations. This finding further illustrates that driver fatigue, especially for large truck drivers, is a problematic issue that needs to be thoroughly investigated so that effective countermeasures could be used. As such, the trucking industry and traffic safety agencies should work more closely together to close the gap and start conversation to identify more effective strategies to monitor fatigue of drivers of large truck by possibly investigating effective outreach campaigns, changing laws, policies, and regulations to enforce large truck drivers to adhere to HOS rules, assessing driver fatigue susceptibility, and improving fatigue training programs for drivers.

3. Non-incapacitating and possible injuries were more likely to be sustained by large truck drivers involved in ROR crashes occurring on horizontal curves. This implies that effective engineering countermeasures should be installed on these locations to reduce the frequency and severity of ROR crashes involving large trucks. This could be achieved by installing curve warning signs, dynamic curve warning system, raised pavement markers, and wider edge lines to alert drivers of an abrupt change in a roadway alignment (Nambisan and Hallmark, 2011).

4. ROR crashes involving large trucks, in which drivers did not wear a seatbelt were found to have a higher probability of sustaining fatal and/or incapacitating injuries. This finding underscores the crucial role that seatbelt plays in saving large truck drivers’ lives. A similar finding has been found in other studies (Anderson and Hernandez, 2017; Chang and Mannering, 1999; Chen and Chen,
2011; Peng and Boyle, 2012). More emphasis on public awareness campaigns could increase seatbelt usage among truck drivers.

5. The origin of drivers played a vital role in injury severity outcomes from ROR crashes. That is, non-Oregonian drivers and those who get their driving licenses from other states experience no injuries. This could be attributed to the driver behavior in the sense that drivers who are not from Oregon may tend to compensate for road network unfamiliarity leading to more cautious driving. This finding is in line with Harb et al. (2008), in which they found that non- Floridian drivers were less likely to be involved in a work-zone crash compared to local drivers.

In summary, this study provides several practical and policy implications that could aid state agencies to improve truck safety. This can be done by a wisdom allocation of limited resources (i.e. labor, time, money) to enhance ROR crashes involving large trucks by utilizing proven countermeasures.

5.3 Research Implications

Transportation agencies strive to prevent roadway crashes and its consequences via safety intervention programs such as deterring alcohol-impaired driving, aggressive driving discouragement, truck safety programs, and reducing driver distractions. The efficiency of those safety programs can be translated into reducing the frequency of roadway crashes and its severities. This, definitely, can lead to significant saving in human lives and social-economic costs. Yet, due to limited funds allocated, prioritizing the safety countermeasures based on their effectiveness in deterring roadway crashes would be used by state agencies. With this in mind, developing advanced econometric methods to identify contributing factors to injury severity of large truck drivers in ROR crashes involving large trucks while addressing the unobserved heterogeneity in the crash data can be used as a basis to aid transportation engineers, trucking industry,
transportation planners, and state agencies in implementing the appropriate safety countermeasures to alleviate ROR crashes.

In terms of study implications, the estimation results of the current research have direct implications on safety policy. First, in relation to seatbelt non-user large truck drivers, the finding of the current study supports the notion that the current policy regarding seatbelt usage should be amended in an attempt to increase seatbelt usage by penalizing drivers violating this policy to deterrent fines and forcing them to enroll in a required defensive driving course. As such, “Click it or Defensive Driving and Ticket” campaign should replace the current campaign “Click it or Ticket” in the state of Oregon. Second, as regards to horizontal curves, a policy maker should implement strategies to reduce the frequency and severity of curve-related ROR crashes involving large trucks to prevent vehicles from leaving their lanes at horizontal curves and minimizing the severities of such crashes if curve-related ROR crashes occur. This could be achieved through using some cost-effective and proven countermeasures such as installing shoulder and centerline rumble strips, enhancing sight distance at horizontal curves by removing or allocating roadside fixed objects such as trees, providing good street lighting, delineation of curves, and deploying advance curve warning systems. Third, the estimation results reveal how driver fatigue increases the probability of non-incapacitating and possible injuries in ROR crashes involving large trucks. As such fatigue management programs from a policy standpoint should be conducted. This could be achieved by mandating trucking industries to conduct fatigue awareness training courses for their drivers, monitoring driver fatigue, restricting drivers working hours, prohibiting of long shifts to reduce fatigue, and installing electronic logging devices on trucks to monitor truck operations and movement.

In summary, this study provides several practical and policy implications that could envision state agencies to improve truck safety. This can be done by a wisdom allocation of limited resources (i.e. labor, time, money) to enhance ROR crashes involving large trucks by utilizing proven countermeasures.
5.4 Practical Applications

The current dissertation contributes, practically, to a better understanding of factors contributing to ROR crashes involving large trucks in the state of Oregon. This understanding can aid Department of Transportations (DOTs) and the trucking industry develop, test and implement effective countermeasures to mitigate such crashes. In terms of econometric models used, such advanced econometric frameworks will not only contribute to the evolution of transportation safety, but eventually can help transportation safety professionals in mitigating future crashes. For instance, by allowing them (the analyst) to account for unobserved factors not currently captured by current methods. In addition, these more advanced econometric approaches can shed light on future data collection strategies to improve injury severity estimation.

5.5 Limitations and Directions for Future Research

In spite of the interesting findings and valuable insight of this dissertation, however, it is not free from limitations. Among these limitations is the unavailability of some important attributes in the Oregon crash data such as pavement conditions, shoulder type, shoulder width, lane width, driving experience, the actual speed of a truck prior to the crash, the degree of curvature for curves, and roadside characteristics. Missing such information can potentially increase the probability of model specification errors. Further, this study primarily focused on ROR crashes involving large trucks in the state of Oregon. As such, the findings of this dissertation are subject to the type of crash data used and the utilized methodological approaches. As a result, applying these findings to other states may lead to temporal and spatial variation.

With respect to future research, some potential room for future improvements should be considered. In terms of methodological approach, developing other methods that address the issue of unobserved heterogeneity would be recommended as a venue for future research. These methods include latent-class models with random parameters
within classes, Markov-switching models, and Markov-switching models with random parameters. Second, future work may consider collecting a comprehensive crash dataset by including the missing variables in this study in the future analyses so that quantifiable models could be developed to capture the factors that contribute to ROR crashes involving large trucks. Fortunately, this is not an impossible task due to the continuing improvements in the data, especially police crash reports. Lastly, different subpopulations such as driver age, gender, roadway classification, number of vehicles involved, and weather conditions could also be investigated in the future studies to better understand factors influencing injury severity resulting from such crashes in each distinct group within a population.
References


Chapter 6: References


Behnood, A., Mannering, F.L., 2015. The temporal stability of factors affecting driver-


Pahukula, J., Hernandez, S., Unnikrishnan, A., 2015. A time of day analysis of crashes
involving large trucks in urban areas. Accident Analysis and Prevention 75, 155–163.


