2013

A study of single and multiple vehicle crashes involving heavy trucks in Iowa

Donald Mathew Cerwick
Iowa State University

A study of single and multiple vehicle crashes involving heavy trucks in Iowa

by

Donald Mathew Cerwick III

A thesis submitted to the graduate faculty
in partial fulfillment of the requirements for the degree of
MASTER OF SCIENCE

Major: Civil Engineering (Transportation Engineering)

Program of Study Committee:
Konstantina Gkritza, Major Professor
Shashi Nambisan
JiangPing Zhou

Iowa State University
Ames, Iowa

2013

Dedication

I would like to thank my parents Don and Sharon, my sister Andria, and my best friends Chase and Callie for supporting and encouraging my decision to abandon my previous career and pursue my graduate degree. It was the best decision I ever made!

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

First, I would like to thank Dr. Konstantina Gkritza for going against her better judgment and accepting the challenge of being my Major Professor. Her patience, support, and encouragement have been immeasurable and I cannot express my gratitude enough.

I would also like to thank my Program of Study Committee Members, Dr. Shashi Nambisan and Dr. JiangPing Zhou for their useful comments and suggestions which shaped my final analysis.

I would also like to extend my gratitude to Mr. Zach Hans for his assistance with the data processing and all the insightful recommendations he provided on a daily basis.

ABSTRACT

In 2010, 16.5 percent of all fatal vehicle crashes in Iowa involved large trucks compared to the national average of 7.8 percent. Only about 16 percent of these fatalities involved the occupants of the heavy vehicles, meaning that a majority of the fatalities in fatal crashes involve non-heavy truck occupants. These statistics demonstrate the severe nature of heavy truck crashes and underscore the serious impact that these crashes can have on the traveling public. These statistics also indicate Iowa may have a disproportionately higher safety risk compared to the nation with respect to heavy truck safety. Several national studies, and a few statewide studies have investigated large truck crashes, however no rigorous analysis of heavy truck crashes has been conducted for the state of Iowa. This thesis uses the most current statewide crash data to perform an in-depth analysis of heavy truck crashes in Iowa. The objective of this study is to investigate and identify the causes, locations, and other factors related to heavy truck crashes in Iowa.

To conduct this study, crash data for the years of 2007-2012 for the state of Iowa were used to develop statistical models for single and multiple vehicle heavy truck crash severity. Single vehicle crashes were modeled using a binary probit model with outcomes of injury (fatal, major, minor, or possible injury) or no injury (property damage only). Multiple vehicle crashes were modeled using a nested logit model with severity outcomes of severe injury (fatal or major injury), minor injury (minor or possible injury), and no injury (property damage only), with the two injury outcomes placed in a nest.

Findings from the two models were both complimentary and contradictory. Both models found older drivers to be associated with more severe injuries. Both models also indicated crashes impacting and damaging the front of both heavy trucks and non-heavy trucks to play a significant role in the severity outcome of the crash. The main disparity of the two models relates to the effect single unit and combination trucks have on crash severity, with combination trucks increasing the probability of a severe injury in the multiple vehicle model and single unit trucks increasing the probability of an injury in single vehicle crashes. Other factors found to be significant in either of the two models relate to the manner of the collision, temporal factors (season, day of week, time of day), vehicle characteristics, roadway characteristics, and environmental factors.

CHAPTER 1 INTRODUCTION

1.1 Research Motivation

In 2010, 16.5 percent of all fatal vehicle crashes in Iowa involved large trucks compared to the national average of 7.8 percent and averages for similar states of 10.3 percent (South Dakota), 19.7 percent (Nebraska), 12.4 percent (Kansas), and 6.6 percent (Missouri) (NHTSA, 2011). In the same year, heavy vehicles represented only 11.8 percent of the VMT in the state of Iowa, indicating heavy vehicle may be overrepresented in fatal crashes (Iowa DOT). Further, between 2006 and 2010 in Iowa, there were on average 74 heavy vehicle involved fatal crashes annually (NHTSA, 2011). Only about 16 percent of these fatalities involved the occupants of the large trucks, meaning that a majority of the fatalities in fatal crashes involve non-heavy truck occupants (NHTSA, 2011). These statistics demonstrate the severe nature of heavy truck crashes and underscores the serious impact that these crashes can have on the traveling public. The statistics presented above also indicate that Iowa may potentially have a disproportionately higher safety risk compared to the rest of the nation and neighboring states (except for Nebraska) with respect to heavy truck safety. Several national studies, and a few statewide studies have investigated large truck crashes, however no rigorous analysis of heavy truck crashes has been conducted for the state of Iowa. This thesis uses the most current statewide crash data to perform an in-depth analysis of heavy truck crashes in Iowa.

1.2 Objectives and Anticipated Results

The goal of this thesis is to investigate the causes, locations, and other factors related to heavy truck crashes in Iowa with a focus on the years from 2007 to 2012. Descriptive analysis, statistical tests, and statistical modeling were used to discover what factors contribute to heavy truck crashes and the corresponding magnitude of the effect of each factor.

Findings of this research will be of interest to multiple parties. Law enforcement agencies will be able to utilize this study’s results to establish enforcement priorities and make determinations on how to best allocate their limited resources to promote safety and reduce crashes. Those with a stake in the freight industry, namely intrastate and interstate carriers, could use the results from this study to better educate fleet managers, drivers, and maintenance personnel, as well as, make changes to when and where to operate their equipment, eventually leading to cost savings and increased productivity through a reduction in crash involvement.

Both state and local planning personnel could also use the results of this study. Lawmakers could use outcomes of this study to assist in the development of laws and regulations in relation to transportation in general or specifically to large trucks and their operations. Those in the transportation financial planning arena may also be able to use the results of this analysis to establish or support funding priorities directed toward improved mobility and safety. Findings in relation to heavy truck crash causes, locations, and demographics could also be utilized by planning personnel to develop targeted educational measures aimed toward the promotion of roadway safety and crash reduction.

Roadway and vehicle engineers may also find the results of this study useful. Identification of what roadway factors leading up to and present at the time of a crash would undoubtedly be of use in future roadway updates and designs. Further, relationships found between vehicular points of contact or the crash's most severe event and the associated severity outcomes could be used to improve and modify vehicle crash attenuation structures and various other vehicular control systems.

1.3 Research Approach

1.3.1 General Information and Definitions

As mentioned previously, this thesis is the first attempt to conduct an in-depth analysis of heavy truck safety for the State of Iowa. Additionally, no extensive work has been conducted on heavy trucks utilizing the same data set used for this study and as such there is no pre-established definition of what a heavy truck is. The vehicles considered for this analysis were carefully selected. A review of similar studies revealed that the definition of what constitutes a heavy truck is quite variable. A heavy truck could be based on the vehicle's weight, the licensure requirements to operate the vehicle, or the vehicle's DOT registration. For this analysis the choice what of constitutes a heavy truck was based largely on complimentary suggestions from members of the Iowa Motor Vehicle Enforcement (Iowa

MVE) and the Federal Motor Carrier Safety Administration (FMCSA). The vehicles suggested and used in the study include both single unit and combination trucks. A sample of the vehicles and categories of vehicles considered can be seen in Figure 1-1. It should be mentioned that a majority of these vehicles, but not all of these vehicles, require a commercial driver's license (CDL) to operate.

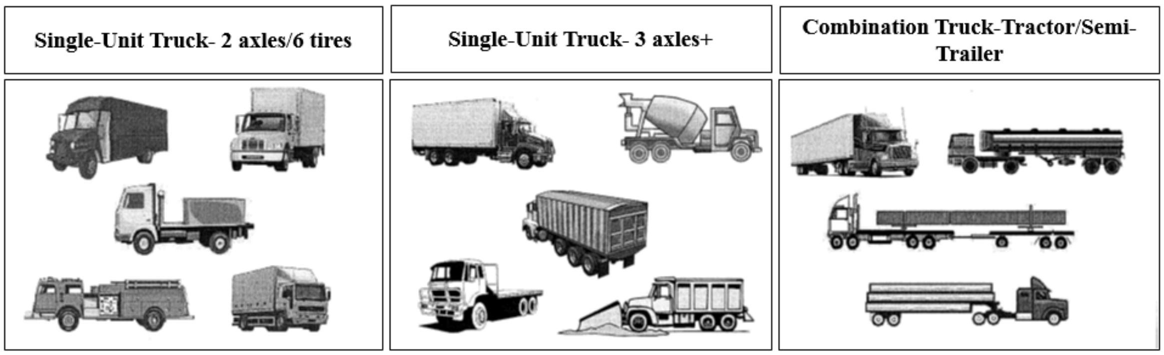


Figure 1-1: Vehicles Considered to be Heavy Trucks

The data set itself and the sources of the data used for this study should also be defined here. The data used for this study, which will be discussed in greater detail later, comes from law enforcement crash reports and includes information on the driver involved, the vehicle involved, the crash location, the time of the crash, the environmental conditions present at the time of the crash, the severity outcome of the crash, and various other factors related to the crash and its possible causes. It should be noted that some of the information populated in the crash reports is subjective and left to the discretion of the officer completing the crash form. All information included in the crash report is populated after the crash has taken place and is based on the observations of trained law enforcement personnel and the information the law enforcement personnel gather from eyewitnesses.

1.3.2 Research Framework

General safety statistics and reports from state and federal sources were first consulted to identify issues and shortfalls related to heavy truck safety. Once the issues were identified, a variety of resources, all with a vested interest in heavy truck and traffic safety, needed to be consulted to establish a suitable data set for analysis, postulations about the causes of heavy truck crashes, and what kinds of analyses would be most helpful to those interested.

Meetings and correspondence between members of Iowa MVE and the FMCSA provided great insight as to what factors have historically been associated with heavy truck crashes in Iowa and what measures have been taken in the past with respect to heavy truck safety and enforcement.

To test the postulations and ensure the results obtained were meaningful, a research methodology needed to be carefully developed. The development of a methodology relied heavily on methodologies and suggestions from past works similar in nature. A multitude of peer-reviewed publications and scholarly reports were critically reviewed for best practices and shortcomings. From the review, a list of potential methodologies was identified and the determination of the methodology to use was made based on the careful consideration of each method's advantages and drawbacks. A more detailed discussion on the selection of a methodology is presented in chapter 4. Outputs from the selected methodology were reviewed for any unsatisfactory or unreasonable results. Questionable findings or violations of assumptions necessary for application of the chosen methodology constituted a review of the present methodology or the selection of an entirely new methodology until an acceptable result was output. A high level look at the framework used for this analysis can be seen in Figure 1-2.

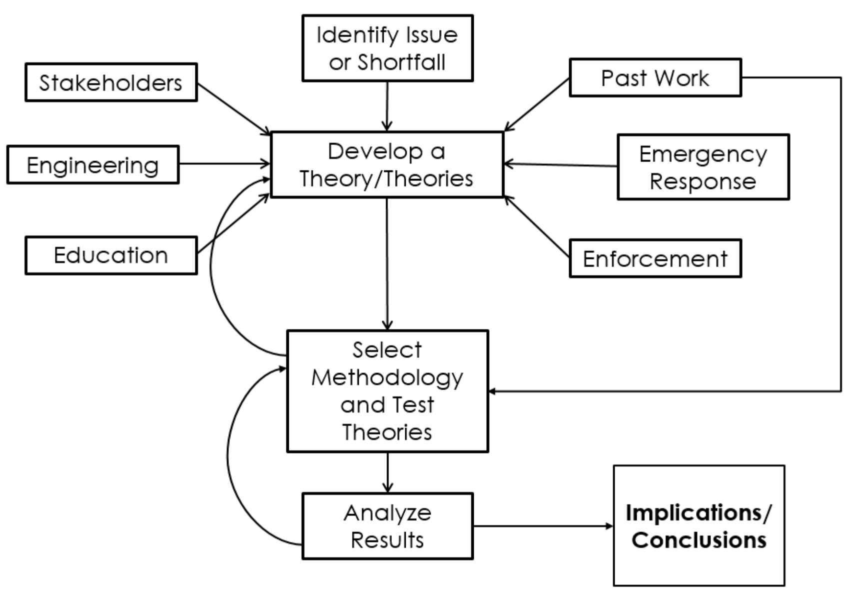


Figure 1-2: Research Framework

1.4 Thesis Outline

This thesis is divided into six chapters.

Chapter 2: *Literature Review* provides an overview of past studies related to vehicle and heavy truck safety. This chapter mainly focuses on various methods of modeling heavy truck crashes and the results of the various studies reviewed. An additional review of possible countermeasures is also included in this chapter.

Chapter 3: *Data Description* provides details of the data set used for analysis. Descriptive statistics, plots of trends, and statistical tests were all utilized to identify variables of interest and visualize the data set used for analysis. Comparisons between the characteristics of the crash data set and the general population are also included in this chapter as a means of better understanding trends observed in the crash data.

Chapter 4: *Methodology* summarizes the methods of analysis utilized in past studies and discusses the benefits and detriments of some of the more commonly used models for performing a crash severity analysis. The characteristics of the discrete outcome models employed for analysis are also discussed. Information on model: specification, interpretation, testing, goodness of fit, and validation of proper functional form is also included in the chapter on methodology.

Chapter 5: *Results* presents the modeling results for both single and multiple vehicle heavy truck crashes. Also included in the results chapter is a discussion of the magnitude, sign, statistical significance, and effect of the variables included in the models as well as a conversation on the overall fit and suitability of the model/s used.

Chapter6: *Conclusions, Limitations, and Recommendations* offers concluding statements on the research conducted. Additionally limitations of the current study are discussed along with suggestions for future research.

CHAPTER 2 LITERATURE REVIEW

2.1 Literature Review Overview

There have been several national-level and state-level studies on commercial motor vehicle (CMV) crash severity. These studies vary in methodology and range from observational/field studies to more rigorous studies involving statistical modeling. From the review of the existing literature, it became apparent that traffic crashes are the result of a complex interaction of numerous factors including driver characteristics, vehicle condition/configuration, environmental characteristics, roadway features/geometrics, and traffic characteristics. Additionally, an analysis of countermeasures aimed at improving commercial motor vehicle safety through changes in roadways, vehicles, and enforcement was conducted and reported. A comprehensive overview of the review findings is presented next.

2.2 Field Studies of Heavy Truck Crash Frequency and Severity

The Federal Motor Carrier Safety Administration's (FMCSA) 2006 Report to Congress on the Large Truck Crash Causation Study (FMCSA, 2006) outlined and identified factors of large truck crashes that need investigation. The study looked at a nationally representative sample of large truck involved fatal and injury crashes in the United States between 2001 and 2003. Vehicles considered to be large trucks included single unit trucks (two and three axles) and combination trucks (truck trailers, tractor trailers). The standard, single tractor trailer configuration, accounted for over 60 percent of the trucks included in the study. From the study it was indicated that trucks were at fault in 55 percent of all crashes (single and multiple vehicle crashes) and 44 percent of all truck/passenger vehicle crashes. The study also noted that driver- related factors accounted for 87-89 percent of the crashes analyzed. The most common factors being traveling too fast for conditions, making an illegal maneuver, legal drug use, unfamiliarity with the roadway, and fatigue. It was noted that fatigue was recorded twice as often for the passenger vehicle driver than for the truck driver. The study also found certain vehicle and roadway characteristics to contribute to large truck crash occurrence, but such factors were far less common than driver related factors. The most common vehicle-associated factor was brake problems and the most common roadway factor was interruptions in traffic flow. The outcome of the study drew no clear conclusions on the causes of large truck crashes, but provided a multitude of guidance that was used in many of the studies discussed within the remainder of this literature review.

A study by Blower and Kostyniuk (2007) used 2001-2005 data from the Michigan Vehicle Crash Files, Trucks Involved in Fatal Accidents file (TIFA), the Motor Carrier Management Information System (MCMIS), and the Michigan FACT file to conduct a descriptive study aimed at identifying the issues that contributed most to commercial vehicle crashes, fatalities, and injuries in the state of Michigan. The result of their analysis indicated that numerous factors, ranging from the driver to the roadway to the vehicle and even the location contributed to severe commercial vehicle crashes. It was found that younger driver crashes were more likely to be coded with hazardous actions such as following too closely or speeding. Younger drivers were also found to be more likely to be involved in backing-up crashes than older drivers. It was also noted that in approximately half of the commercial motor vehicle (CMV) crashes, the hazardous action contributing to the crash was coded for the driver of the other vehicle (non CMV). It was also found that fatigue-related CMV crashes tended to be rear end and single vehicle crashes, with most crashes occurring at night on interstate roads between midnight and 6a.m. Additionally, when all levels of severity were considered, angle crashes, rear end crashes, head on crashes, same direction sideswipes, and single vehicle crashes tended to, in the order presented, contribute the most to CMV crash costs and harm to society. Vehicle defects and inspection violations were also analyzed by Blower and Kostyniuk. It was noted that lighting and brake violations were the most frequent violations in CMV inspections with both smaller fleet carriers and intrastate carriers tending to have higher violation rates in their inspections. It was also observed that intrastate carriers had more serious violations then did interstate carriers. The results provide no clear solution, but suggest strategies to improve commercial vehicle safety will have to work on many fronts, ranging programs to improve the conditions of the vehicles themselves, to programs educating all drivers sharing the road.

2.3 Statistical Modeling of Crash Frequency, Occurrence, and Severity

2.3.1 Crash Frequency and Occurrence Models

Multiple studies have investigated what driver factors contribute to heavy vehicle crashes. A study by Cantor et al. (2009) applied a Poisson regression model on national commercial driver's license (CDL) and crash data to investigate the relationship between driver characteristics and heavy vehicle crashes. The results showed that poor driver safety performance (expressed as number of previous crashes), driver out of service violations, driver body mass index, driver gender, driver age, and past employment were significant characteristics in the prediction of heavy vehicle crash rates. In particular, the model estimated males and drivers under 25 years old to be associated with higher crash rates.

Another study by Park and Jovanis (2010) looked at the effect hours of service and schedules had on the probability of a crash occurring (crash odds). For their study, they collected detailed crash and driving schedule data from three national companies, with varying operations, for a total of 231 crashes. Their primary method of analysis utilized time-dependent logistic regression models to assess the relationship between hours of service/schedule and crash risk. From their models, it was found that the odds of a crash occurring was, indeed, associated with the hours of driving, with particular emphasis placed on times after the sixth hour of driving. With respect to the first hour of driving, the odds of a crash occurring increased by 56 percent after the 6^{th} hour and more that 200 percent after the 10^{th} and 11^{th} hours. The study also found that off duty times of more than 46 hours were associated with an increase in crash risk. These findings are of great interest and provide ample guidance, however these findings were obtained based on a limited sample size.

A similar study conducted by the U.S. Department of Transportation and the Virginia Tech Transportation Institute (Barr et al., 2011) analyzed driver drowsiness to assess the impact that drowsiness had on commercial motor vehicle driving performance. Their research objectives included characterizing the occurrence of drowsiness and its cause(s); exploring the effects of drowsiness on safe driving performance; and identifying relationships between drowsiness, distraction, and performance. Data were collected as part of a naturalistic field study. Cameras filmed drivers and lane position. A total of 908 hours of video footage was collected and then processed. Drowsiness events observed from the videos were then documented, described, and entered into a data set. Analysis of Variance (ANOVA) tests, stepwise linear regression, and logistic regression were then used to analyze the collected data. Generally, all three of the data analysis methods produced consistent results. Each analysis method showed evidence of a strong correlation between drowsiness and the time of day, with early morning time periods between 6am and 9am being particularly problematic. The opposite finding was observed between the hours of 12pm and 3pm when drivers appeared to be alert. These findings indicate that drivers may not be fully refreshed or awake in the early hours of their work shift and special precautions during these hours may be of great aid to the drivers and the traveling public. Drowsiness was also found to be related to age and experience. Younger drivers in the 19-25 year old age group were found to be nine times more likely to be classified in the “high fatigue” group of drivers. Similarly, inexperienced drivers with less than one year of commercial driving experience were found to be seven times more likely to be grouped in the “high fatigue” category. The results of this study provided some interesting results with important implications especially related to younger and inexperienced drivers.

A study by Blower et al. (2010) used the data and findings from the Large Truck Crash Causation study to examine the relationship between vehicle condition and crash involvement in more detail. More specifically the study attempted to test two different hypotheses. The first hypothesis tested was that trucks with defects and out of service (OOS) conditions are statistically more likely to be in the role of actuating a crash then trucks with no defects. The second hypothesis tested was that defects in specific systems are associated with crash roles in which those systems are paramount in crash avoidance (a physical mechanism links the vehicle defect to the crash). To test these hypotheses, multiple logistic regression models were developed to show if any statistical association was present. From the models it was found that the critical reason for the crash was mostly associated with driver factors and less likely due to a mechanical defect. Among all mechanical systems, only brakes were shown to be significantly statistically related to the crash cause. More specifically brake adjustment was found to be most significant mechanical defect associated with the cause of a crash. The results of this study, though limited, do identify two key aspects: first, drivers are clearly a critical factor in truck crashes; second, mechanical conditions do, to a lesser extent than drivers, also play a role in truck crashes with a key emphasis placed on the brake systems.

A study conducted by Giuliano et al. (2009) used both descriptive statistics and statistical modeling to analyze the factors and trends associated with commercial motor vehicle crashes in the state of California. From the descriptive investigation, it was observed that the fewest crashes occur in the winter and early spring (January, February, and April) and crashes peak during the late summer and early fall (August, September, and October). It was also observed that few crashes occur during the late night and early morning, but crash occurrences tended to rise throughout the morning, peak in the early afternoon, and then dramatically reduce in occurrence after 6 PM. Additionally the researches also noticed a crash pattern by day of the week. The data indicated that crashes tended to be most frequent on Tuesday and Friday and minimal over the weekend. In an effort to gain further insight into the crash phenomenon both a Poisson and a Weighted Least Squares (WLS) model were developed based on county level data. Both models contained the same variables and reported similar findings. From the models it was interpreted that precipitation, the percentage of elderly residents, and the percentage of foreign born residents were all strongly and significantly related to an increase in the number of crashes. One surprising result of the models was the indication that heavily urbanized areas are actually less dangerous for trucks than more rural areas. The only variables the two models reported different signs for were variables related to road usage and the percentage of young residents. The WLS model indicated that increases in road usage and the proportion of younger residents in the population would lead to an increase in crash frequency, but the opposite relationship was expressed in the Poisson model. However, no conclusions were drawn as to whether one model was preferred to the other.

2.3.2 Crash Severity Models

Binary Models

A study published by the National Center for Statistics and Analysis (Moonesinghe et al., 2003) looked at how the environment and the characteristics of the vehicle impact a truck's propensity to roll over or jackknife in single-vehicle collisions. To conduct the analysis, data from the Trucks Involved in Fatal Accidents survey (TIFA) was used. From the TIFA data a binary logit model was developed to estimate the probability of a large truck rollover or jackknife. The model's results suggested that a speed limit of 55mph or higher, poor weather, and a curved road all substantially increases the odds of both a rollover or a jackknife occurring. Additionally, it was found that the odds of a rollover increased with increasing the weight of the large truck and cargo, but the odds of a jackknife actually decreased with increasing the weight of the large truck and cargo. However, opposite results were found for increases in truck length. These results are specific to just rollover and jackknife occurrences, but the findings and methodology are still of use in analyzing heavy vehicle crashes.

Bham et al. (2012) used a multinomial logit (MNL) model to examine the differences in crash contributing factors for six collision types, and a binary logit model to identify factors that contribute to crash injury severity (severe and non-severe crashes) for motor vehicles in Arkansas. The multinomial model's estimation results suggested that the risk of a multi-vehicle crash was higher during weekdays while the risk of a single vehicle collision was higher over the weekend. It was also deduced that single vehicle collisions were significantly associated with nighttime and wet conditions. The binary logit model of injury severity showed that drivers who did not wear a seatbelt and those under the influence of alcohol were more prone to severe crashes. The binary model also indicated that roadway grades and the presence of curves also increased the severity of crashes. Another notable finding from the binary severity model was that the severity of crashes actually declined under wet roadway conditions, which is likely due to drivers being more attentive and cautious under such conditions.

Ordered Models

Lemp et al. (2011) used both and ordered probit and heteroskedastic ordered probit (HOP) model to study the impact of vehicle, occupant, driver, and environmental characteristics on the injury severity outcome of large truck crashes. Data used for this study came from the United States' Large Truck Crash Causation Study, General Estimates System, and Vehicle Inventory and Use Survey. Factors, found by both models, to increase the severity outcome of a large truck crash include multiple vehicle crashes, multiple occupant vehicles, crashed involving more than one truck, and crashes occurring under dark lighting conditions. Generally, both models produced consistent results, however it was determined that the more flexible HOP model performed significantly better.

A study by Abdel-Aty (2003) used multiple ordered probit models to investigate motor vehicle crash severity for roadway sections, signalized intersections, and toll plazas in Florida. The four levels of severity incorporated in the models were no injury, possible injury, evident injury, and severe/fatal injury crashes. Several factors were common across all the models and those factors were driver age, gender, seatbelt use, vehicle type, point of impact, and speed. From the models developed it was found that elderly drivers, those not wearing seatbelts, and male drivers all have a higher probability of severe injuries. The modeling results also highlight that other factors related to the location of the crash contribute to higher severity levels. Such location specific factors associated with high severity include characteristics such as roadway curves, dark lighting conditions, and rural areas. Other modeling approaches such as multinomial logit models and nested logit models were attempted, but the results of these models were rather poor in comparison to the ordered probit model discussed previously.

A different study by O'Donnell and Connor (1996) utilized both an ordered probit and an ordered logit model to model the relationship between crash severity and the attributes of motor vehicle users in New South Wales, Australia . The study found that higher speeds, high blood alcohol content, older vehicles, and older drivers were highly linked to greater crash severity. It was also found that the vehicle type and vehicle manufacturers (brand) were also significant determinants of crash severity.

A similar study for heavy vehicles conducted by Kockelman and Kweon (2002) also employed an ordered probit model to estimate crash severity. From the model's results a variety of implications could be drawn. It was determined that the manner of collision, number of vehicles involved, driver gender, vehicle type, and alcohol use all played a significant role in crash severity. The results also corresponded well with the works discussed earlier by O'Donnell and Connor on motor vehicle users.

Unordered Models

Environmental factors such as the weather, the type of roadway, and the area surrounding a roadway also contribute to heavy vehicle crashes and crash severities. In one study conducted by Khorashadi et al. (2005) heavy vehicle crash severity was examined in urban and rural areas. This study used a multinomial logit (MNL) model to model four outcomes of heavy vehicle crash severity (no injury, complaint of pain, visible injury, severe/fatal injury) in urban and rural conditions, with severe crashes being more prevalent in rural areas. Their study found some striking differences between the two area types and their respective models. Most notable was that the different models contained different variables. Multiple variables found to be significant in the urban model, turned out to be insignificant in the rural model and vice versa. Additionally, variables shared by both models typically possessed signs of different magnitude and impact. These findings underscore the difference between urban and rural large truck crash severities and suggest that complex interactions between driver and other measurable environmental factors are playing a significant role in the demands placed on the driver in rural versus urban areas.

Cheng and Mannering (1999) used two nested logit models to determine the influence that certain factors have on the injury severity outcome of both truck and non-truck involved accidents. The nest structure can be seen in Figure 2-1. The data used for the project was for King County in Washington State and included information regarding injury, weather, alcohol use, restraint use, roadway conditions, and factors contributing to the accident. Both the truck and non-truck models were compared for similarities and differences. One variable that was unique to impact trucks was a variable for speeds of 55 mph. The speed variable increased the likelihood of possible injury and injury/fatality outcomes, but was found to be insignificant in the non-truck model, highlighting the critical relationship between speed and truck crash severity. Other variables found to only be significant in the truck model included variables for left or right turns and rear end crashes. To supplement the comparison between trucks and non-trucks, elasticity's were computed and compared. From the elasticity analysis it was found that the variables common to both models generally had a much larger impact on the outcome of the truck model which underscores the great importance and potential impact of truck safety countermeasures.

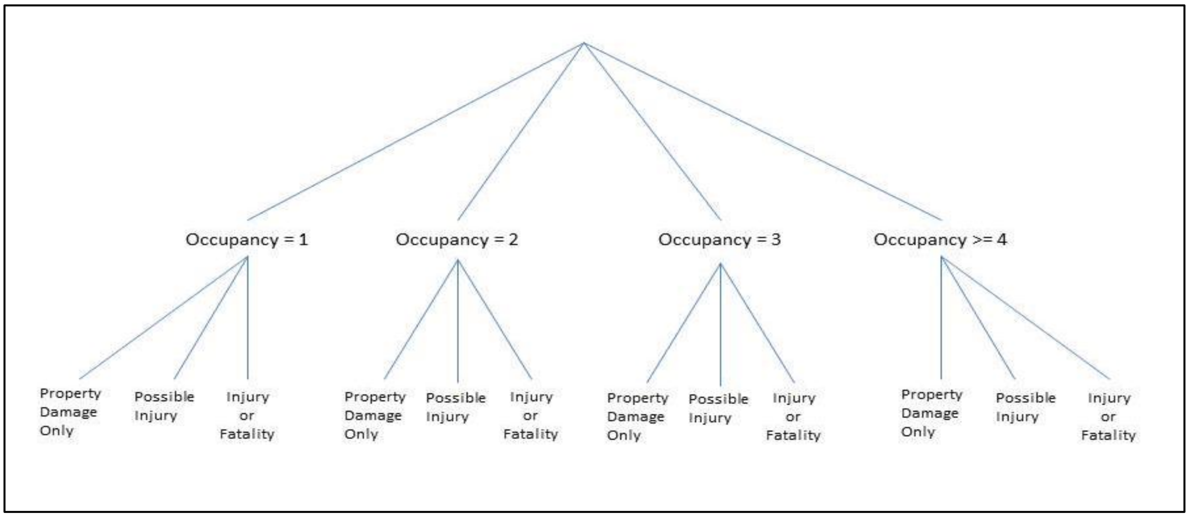


Figure 2-1: Proposed Nested Logit Model Structure Source: Chang and Mannering, 1999

Other discrete outcome models such as latent class logit models (LCL) have also shown to be effective. A study by Xie (Xie et al., 2012) examined motor vehicle driver severity in rural single vehicle collisions. For this study researchers created both an MNL model and LCL model to analyze the same data set. Both models were run with the same 31 explanatory variables that included information on traffic, roadway geometry, driver characteristics, vehicle characteristics, and environmental characteristics. Variables for driver age, alcohol use, lighting conditions, speed, and ethnicity were all significant variables in the determination of crash severity in both models. It was also noted that the variables in both models were consistent in both the signs and trends of their marginal effects. To further compare the two model types, a prediction experiment was conducted to evaluate the goodness of fit of both models. From the experiment it was determined that the LCL model generated a satisfactory fit and prediction ability, and when compared to the MNL model, the LCL model improved prediction accuracy by 37 percent. This result is encouraging, but the authors suggest additional testing be performed before a conclusion can be drawn on the use of LCL models over MNL models.

Non-parametric modeling methods have also been used to establish a relationship between injury/severity outcome and driver, vehicle, environmental, and roadway conditions. A study conducted by Chang and Chien (Chang and Chien, 2012) used a non-parametric

Classification and Regression Tree (CART) model to investigate the factors associated with truck involved crash severity. The benefit of the CART model is that it is not susceptible to the assumption violations and the associated erroneous estimation results that can plague parametric regression models such as MNL models and ordered regression models. The results of the CART model were comparable to many past studies and, for the most part, reinforce many of the findings already discussed. However, despite the misspecification advantage, the CART model was limited in usefulness. Elasticity's and marginal effects for each injury outcome cannot be calculated from a CART model's output and as such CART models are not able to fully and correctly evaluate the relative impact of each variable in the model.

In summary, a review of the literature clearly shows that statistical modeling is a proven tool capable of analyzing vehicular crashes and the factors that contribute to the crashes themselves. However, once contributing factors are identified, the next challenge becomes implementing practices that can favorably alter these factors. Practices targeted toward improving roadways, vehicles, and enforcement have been developed and show promise at reducing both the occurrence and severity outcome of crashes. An overview of these potential countermeasures follows.

2.4 Countermeasures

2.4.1 Roadway Improvements

One strategy for improving truck safety involves making changes to the existing roadway and roadway regulations. In a study conducted by Harwood et al. (2003) researchers used findings from interviews and literature reviews to analyze the interaction between commercial trucks and busses with highway features. The researchers found that traffic control devices and traffic regulations play a significant role in the safe movement of heavy vehicles. In particular, the researches mentioned safety benefits are capable through the use of differential speed limits, lane use restrictions, exclusive lanes, and modified signal timing. The researchers also noted that the increased use of intelligent transportations systems (ITS) has also been of great benefit to improvements in heavy truck safety. Such ITS systems mentioned were downgrade warning signs, dynamic curve warning systems, and improved weigh stations.

A different report by McMurty et al. (2007) identified some additional roadway design and operations problem areas. Truck's high centers of gravity, longer braking distances, and articulation all contribute to trucks having an increased rollover risk at curves, particularly curves on exit ramps. One countermeasure suggested was truck specific warnings/advisory speeds (both before and during the curve) that incorporate dynamic signing. Vehicles at risk are identified by sensors and dynamic signage is then used to notify the drivers of the impending danger with enough time for corrective measures to be taken. In addition to curves, work zones also present an increased safety risk for heavy vehicles. Some possible work zone countermeasures to consider include rumble stripes, highway advisory radio, and queue detection and warning systems. As with many new technologies there is little work to draw conclusions on effectiveness of any of the improvements mentioned, but none the less there are a multitude of countermeasures available for consideration.

Potter et al. (2013) analyzed heavy truck crashes in urban areas and identified multiple ITS technologies that could potentially decrease the occurrence of heavy truck collisions. From crash data it was noticed that a majority of heavy truck crashes in urban areas were rear end crashes taking place at intersections. Intersections of interest were then selected and site investigations were conducted to indicate potential causes and identify practical ITS solutions. Commonly reported infrastructure ITS improvements included:

- Activated warning signs for queuing and end of green
- Intersection collision avoidance systems using short range radio
- 'Dilemma zone' activated clearance time extension
- Various other vehicle to infrastructure communication systems (speed, rail, clearance, etc.)

2.4.2 Vehicle Improvements

Technological improvements to vehicles have the ability to influence heavy vehicle safety in two ways:

1. Improve the performance of the vehicle (avoid or survive crashes better)

2. Improve the performance of the driver

A report by Blower and Woodruff (2012) outlines an emerging set of new technologies available to help a driver control their vehicle. One technology under development for large trucks is electronic stability control (ESC). ESC is a technology that helps drivers maintain control and prevent a rollover of the vehicle should the driver lose lateral control and begin to roll. In an effort to reduce rear end collisions, both forward collision warning systems (FCW) and collision mitigation braking systems (CMB) are also being considered for use in large trucks. If a driver fails to react to a collision both systems work to alert the driver in an attempt to avoid the collision. The CMB system will actually apply the brakes without input from the driver in an effort to reduce the severity of the crash should the not respond to the FCW system. Another system mentioned was the lane departure warning system (LDW). LDW systems alert a driver should the vehicle inadvertently leave the lane of travel. LDW systems are believed to have the ability to reduce sideswipe crashes as well as reduce crashes resulting from drowsy drivers. In addition to new technologies, improvement of some existing technologies also shows promise. Underride guards presently equipped on trucks in the United States are not strong or low enough to be effective and as such, it is suggested that more work be done with respect to new improvements and regulations relating to current underride prevention systems.

Perrin et al. (2007) discussed many other technological improvements on the horizon to improve heavy vehicle safety. One technology currently under review is the use of electronically controlled braking systems (ECBS). ECBS systems control a vehicle’s brakes electronically rather than pneumatically. Electronic control of the brakes provides for better response, more precise control, and a better platform to introduce the ESC, FCW, and CMB systems mentioned in the previous report. Other improvements discussed include monitoring the driver and the driver’s behaviors. Most of these systems are conceptual at this point, but the idea is to, one provide the driver feedback if the driver presents a risky behavior (drowsiness, speeding, tailgating, etc.), and two to monitor the driver’s hours of service and tendencies in an effort to reduce unsafe behaviors. Preliminary studies in Belgium and the Netherlands showed such systems were capable of reducing crashes by 20 percent, but the issue of intrusion of privacy is a large hurdle to overcome before such technologies are considered for widespread use. Another conceptual technology being considered is the use of wireless communications to support vehicle-to-vehicle and vehicle-to- infrastructure communications in an effort to heighten driver awareness. Details of the possible applications are provided in Table 2-1.

Table 2-1: Examples of Applications of Vehicle Wireless Communications Source: TRC (May 2007)

Public Safety Applications	Private Sector Applications
Vehicle-to-Vehicle Approaching emergency vehicle (warning) Cooperative collision warning Cooperative adaptive cruise control	*All Vehicles* Access control Onboard diagnostic data Repair-service record Vehicle ECU program updates Enhance route planning and guidance
Vehicle-to-Infrastructure Road condition warning Low bridge warning Toll collection Traffic information Green light- optimal speed advisory	*CMVs* Automated vehicle safety inspections Border clearance information (credentialing) Electronic manifests (hazmat) Unique CVO fleet management applications

Other vehicle improvements mentioned were focused on surviving the crash and protecting the occupants. Many of the technologies discussed for the occupants of the large trucks already exist widely. Many trucks are already equipped with seatbelts and front impact air bags and years of testing has shown both of these mechanisms, when used in conjunction, to be rather effective. The use of side impact airbags is rather new however, but shows promise. Studies in Europe have shown side airbags to be a rather effective means in the prevention of ejection and vehicle rollover.

Further improvements discussed were focused on protecting those in the other, light vehicle(s) involved in the collision with the large truck(s). Such technologies under consideration include front underride prevention improvements (also mention by Blower and Woodrooffe), crash-attenuating front structures, and deflecting front structures. Measures taken to improve front underride are rather simple and include modifying existing frontal structures or creating new frontal structures for trucks that are low enough to ensure the truck’s structure engages the crash absorbing mechanism of the light vehicle. Another means of improving the crash outcome of a collision with a heavy vehicle involve the dissipation of collision energy either through crash attenuation structures or energy deflecting structures. Crash attenuation structures dissipate crash energy by allowing the heavy vehicle to crush, collapse, and absorb a crash's energy and thus reduce the severity of the injuries sustained by the humans involved in the crash. Energy deflection, on the contrary, uses structures that manage a collision's energy by deflecting the impacting vehicle through the use of properly designed truck structures. Deflecting a crash's energy reduces the collision energy absorbed by the light vehicle which reduces the resulting injury outcomes, but does increase the possibility of a secondary collision. Many of these proposed systems or structures are theoretical, and development and testing is necessary before any definitive conclusions are drawn.

2.4.3 Enforcement

Another alternative counter measure involves modifying enforcement practices. A study by Strathman et al. (2010) looked to identify program strategies and practices that could potentially be implemented by the Oregon Department of Transportation Motor Carrier Transportation Division in an effort to reduce commercial motor vehicle crashes. To conduct their study, a cluster analysis was implemented to establish peer states with geographic, development, travel, and safety enforcement conditions similar to those found in Oregon. Once peer states were established, structured interviews of each state's Motor Carrier Safety Assistance Program representative were conducted. The states included in the study were Oregon, Colorado, Michigan, Minnesota, Nevada, Washington, Kentucky, and Florida. From the peer interviews a multitude of suggestions were compiled and reported. Though protocols for conducting driver and vehicle inspections are fixed, the interviews did offer some tactics that benefit the effectiveness of inspection activities and they are:

- Having troopers prepare their own regional safety plans
- Placing special enforcement in places where there are no inspection/weigh stations
- Increasing the number of inspectors by using the private sector (e.g., truck repair businesses)

- Using aircraft to spot trucks attempting to bypass stops.

The interviews also supplied additional useful tactics with respect to traffic enforcement practices some of the findings are listed below:

- Joining top performing troopers with inspectors
- Targeting high-risk highway segments
- Using data tools to identify at risk drivers
- Patrolling in unmarked vehicles to identify unsafe automobile drivers around commercial vehicles

Additionally the interviews also revealed various tactics to improve the overall effectiveness of compliance reviews and they are:

- Extending compliance reviews to intrastate carriers
- Maintaining the training of inspectors
- Focusing on "at risk" carriers identified by the Federal Motor Carrier Safety Administration

Relocating enforcement efforts also has the potential to impact road safety. Huges (2000) conducted a study in North Carolina to evaluate a change in enforcement practices and a reallocation of efforts. Between the years of 1998 and 1999 the North Carolina Department of Transportation identified 21 counties as having the most truck involved crashes and as such reallocated and increased CMV enforcement in those 21 targeted counties. The increased CMV enforcement consisted of an increase in roadside inspections, an increase in driver and vehicle out of service violations, an increase in CDL citations, and an increase in public education efforts. The product of these combined efforts produced a 17.7 percent reduction in fatal truck involved crashes for the 21 county area and a 5 percent decrease in truck involved crashes statewide between the years of 1998 and 1999. Counties outside the 21 target counties actually saw a 7.6 percent increase in heavy vehicle involved fatal crashes which highlights the resource dependent nature of CMV enforcement practices and underscores a need for improvements geared toward offsetting manpower and personnel limitations. The study suggests improvements through a systematic reallocation of enforcement efforts is possible, however other methods of improvement should also be considered in the future to ensure available resources are optimally utilized.

McCartt et al. (2007) offered even more suggestions for advancing enforcement techniques. For the most part the suggestions presented focused on compliance programs and a select list of those suggestions is presented below.

- Identifying and focusing on problematic carriers and drivers with relatively poor safety records
- Building databases to support problem identification
- Increasing oversight of new drivers and carriers
- Electronic screening bypass systems that allow qualifying carriers, vehicles, and drivers to bypass weigh stations, port-of-entry facilities, and roadside inspections
- Automated vehicle performance monitoring (i.e., brakes, tires)

A related study by Lucke (1999) used a team of federal, state and industry representatives to survey and assesses the effectiveness and uniformity of roadside vehicle inspections in the U.S. Site visits took place in seven states: Illinois, Arizona, California, Tennessee, Connecticut, Minnesota, and West Virginia. From these site visits observations were reported and best practices were then identified by the project team. Overall the team found that a majority of the inspections observed to be uniformly conducted from state to state and some of the best practices the team found were:

- Use of an inspector evaluation process that focuses on the quality rather than quantity of inspections.
- Working with seasonal carriers during their off season to inspect their vehicles thoroughly
- More outreach programs to make both the commercial vehicle industry and the general public more aware of commercial vehicle safety.
- Further utilization of technology to permit both the entry and access to real-time commercial vehicle information.
- Requiring driver placed out of service to sign a form that explains the penalties of an out of service order and that they are aware of these penalties.

The best practices identified by Lucke (1999), though broadly detailed, do offer areas for enforcement agencies to focus on and possibly re-evaluate their current practices. This concludes the discussion on countermeasures. A summary and synthesis of all the findings presented throughout the literature review follows.

2.5 Literature Review Summary

Traffic crashes are the result of a complex interaction of numerous factors. One pattern consistently noticed in a review of past studies was that factors relating to the drivers of both large trucks and other vehicles appear to play a disproportionally large role in crash occurrence. Of all the driver factors considered, age, experience, and behavior (speeding, following too closely, etc.) tended to be the most common and most statistically significant factors. Other variables such as gender, physical condition, and ethnicity, though pertinent in some studies, gave mixed and varying results.

Location, environmental, and mechanical factors appear to also contribute to crash occurrence, but to a much lesser extent than driver-related variables. Numerous studies indicated lighting and brake defects to be common mechanical defects on large trucks, with brake defects actually showing a significant correlation to crash occurrence. Other vehicle factors noted to be significant by other studies include vehicle age, load characteristics (weight and length), and carrier type (small/large, interstate/intrastate, long haul/short haul).

Significant spatial and temporal factors were also revealed by past works. Severe heavy vehicle crashes were found to be more likely to occur in rural areas, at night/dark light conditions, at early times of the day, during peak traffic hours, and on curves. Precipitation, though likely to increase crash frequency, was not found to be associated with severe crashes. This finding is likely attributed to drivers being more cautious during adverse weather conditions.

This chapter also discussed the current and future countermeasures the transportation industry is considering or should consider implementing to improve heavy truck safety. Countermeasures mentioned relate to improving driver performance, vehicle performance, roadway ease of use, and enforcement techniques. A majority of the improvements for drivers focused on identifying drowsiness, improving reaction time, and monitoring driving schedules. Improvements to vehicles were concentrated mostly on improving a vehicle's stability and braking efficiency. Other suggestions were directed toward adaptations of enforcement methods and were rather ubiquitous. Some improvement measures suggested were targeted enforcement, mandated preventive maintenance programs, strengthened CDL programs, and increased campaigns to broaden the public's understanding of the hazards associated with heavy vehicles in the traffic stream.

This concludes the discussion on the literature reviewed for this thesis. A brief summary of the methodologies used in many of the studies discussed here is included in the methodology chapter. The next chapter discusses the data set used for this study. The selection of a proper methodology is largely dependent on the phenomenon of interest and the data available and as such a description of the data is presented prior to the methodology chapter.

CHAPTER 3 DATA DESCRIPTION

3.1 Data Overview

Heavy truck crash data were obtained through the Iowa Traffic Safety Data Service (ITSDS) at the Institute for Transportation (InTrans) at Iowa State University. The data are a collection of crash reports completed by state and local law enforcement agencies that are aggregated by the Iowa DOT before becoming available at the ITSDS. The crash data consists of crash, vehicle, driver, and passenger-level characteristics of all vehicles involved in reported fatal, major injury, minor injury, possible injury, and property damage only (PDO) crashes in Iowa between the years of 2002-2012. To gain a better understanding of the current nature of heavy truck crashes in Iowa, it was desired to use the most recent data available, however the 2012 data, in particular, were recent enough that imperfections and missing information were of concern. In an effort to balance the effect of these possible imperfections a six-year analysis period (2007-2012) was chosen over the more traditional five-year analysis period. Table 3-1 provides a comprehensive overview of the crash data by number of vehicles involved (single versus multiple vehicle crash). The remainder of this chapter describes in detail and highlights most of the information shown in Table 3-1.

3.2 Heavy Truck Crash Distribution

Table 3-2 shows that the majority of the crashes analyzed involved a standard semi/tractor trailer combination truck while single unit trucks accounted for less than 35 percent the heavy trucks analyzed.

A geographic information systems software, Arc Map 10, was used to extract all relevant data. All crashes and all vehicles involved in a crash with a heavy truck, as identified in Table 3-2, between the years of 2007-2012 were extracted for a total of 23,538 crashes involving 25,003 heavy trucks and 18,414 other vehicles. The distribution of the other vehicles involved in a crash with a heavy truck can be seen in Table 3-3. Over 96 percent of the non-heavy truck vehicles, in a collision involving a heavy truck, involve some type of a small passenger vehicle, with more than half of the collisions involving a passenger car.

Table 3-1: Summary Statics of Select Variables for Multiple and Single Vehicle Heavy Truck Crashes, 2007-2012

Variables	Multiple Vehicle Mean (Standard Deviation) or Percentage	Single Vehicle Mean (Standard Deviation) or Percentage
Crash Specific Characteristics		
Crash Severity		
Fatal and Major/Minor and Possible/PDO	5.54/22.2/72.3	3.86/20.3/75.8
Fatal, Major, Minor, Possible/PDO	27.7/72.3	23.2/76.8
Manner of Collision		
Non-Collision/Rear-end/Broadside/Head-on/Sideswipe(same direction)/Other/Not Reported/Unknown	2.33/30.0/18.4/3.31/33.1/11.0/1.62/0.23	93.6/0.00/0.05/0.37/0.00/0.26/2.36/3.36
Number of Vehicles per Crash		
Two Vehicles/Three or More Vehicles	91.2/8.82	-
Sequence of Events (1st Event)		
Ran off Road/Crossed Centerline/Rollover/Jackknife/Collision with Pedestrian/Collision with Vehicle/Collision with Other non-Fixed Object/ Collision With Fixed Object/Miscellaneous Event/Other/Not Reported/Unknown	1.14/2.46/0.04/0.21/0.01/82.7/0.49/0.30/0.21/8.79/3.33/0.32	36.1/2.11/4.54/4.17/0.00/0.00/3.88/17.2/0.22/28.8/2.95/0.35
Most Harmful Event		
Ran off Road/Crossed Centerline/Rollover/Jackknife/Collision with Pedestrian/Collision with Vehicle/Collision with Other non-Fixed Object/Animal/ Collision With Fixed Object/Miscellaneous Event/Other/Not Reported/Unknown	0.35/0.82/0.62/0.28/0.03/88.8/0.79//0.08/0.83/0.45/2.56/4.17/0.22	3.92/0.16/29.9/9.78/0.00/0.60/1.87/8.82/33.8/0.66/6.48/3.90/0.18
Vehicle in Collision With a Heavy Truck*		
Heavy Truck/Passenger Vehicle/Van/SUV/Light Truck/Other Vehicle Type	14.6/51.4/8.99/13.2/16.7/2.24	-
Time and Location Characteristics		
Month		
Jan/Feb/Mar/Apr/May/Jun/Jul/Aug/Sep/Oct/Nov/Dec	12.1/10.6/6.64/6.36/6.90/7.55/7.20/7.85/7.46/8.30/6.91/12.2	9.83/10.1/6.52/6.76/7.32/7.78/7.41/7.05/6.99/9.34/9.94/11.0
Season		
Winter/Spring/Summer/Fall	34.8/19.9/22.6/22.7	30.9/20.6/22.2/26.3
Day of Week		
Sun/Mon/Tue/Wed/Thu/Fri/Sat	4.88/17.3/18.6/16.9/17.3/17.5/7.55	
Time of Day		
1:00-1:59/2:00-3:59/4:00-5:59/6:00-7:59/8:00-9:59/10:00-11:59/12:00-13:59/ 14:00-15:59/16:00-17:59/18:00-19:59/20:00-21:59/22:00-23:59/Not Reported	1.64/1.22/2.18/8.95/15.2/15.5/15.3/16.2/11.8/5.91/3.61/2.47/0.03	5.31/5.38/7.09/9.85/12.2/12.0/11.0/10.6/9.32/6.37/6.19/4.69/0.04
Location		
Urban/Rural	59.9/40.1	34.9/65.1

*Indicates indicator variables established by relating crash level information to the vehicle level. This relate often results in a many to one relationship. Values may not add to 100% due to the possibility of a many to one relationship.

Table 3-1: Summary Statics of Select Variables for Multiple and Single Vehicle Heavy Truck Crashes, 2007-2012 Continued

Variables	Multiple Vehicle Mean (Standard Deviation) or Percentage	Single Vehicle Mean (Standard Deviation) or Percentage
Roadway Characteristics		
Speed Limit		
5/10/15/20/25/30/35/40/45/50/55/60/65/Not Reported	0.17/0.32/0.39/1.08/17.2/6.51/13.0/1.54/6.89/1.84/23.9 /0.57/11.3/11.3/4.03	0.20/0.42/0.42/0.79/11.3/2.86/6.17/0.44/3.70/3.22/30.2/ 0.27/12.9/19.3/8.17
Road Classification		
Interstate/US Route/IARoute/Secondary/Municipal/ Institutional/Unknown	25.9/24.0/12.5/10.0/26.7/0.14/0.73	32.6/21.4/10.8/20.9/12.9/0.05/1.34
Road Type		
Four-way Intersection/Other intersection/Non-Intersection/Not Reported/Unknown	23.6/16.6/59.4/0.33/0.09	7.93/13.4/72.9/5.71/0.07
Roadway Contributing Circumstances		
None/Surface Condition/Work zone/Other/Not Reported/Unknown	75.9/15.8/3.33/2.34/2.15/0.48	67.4/18.6/2.38/3.89/6.85/0.88
Vehicle Characteristics		
Heavy Truck Age		
Continuous	7.59 (7.34)	7.28 (7.55)
Heavy Truck Type		
Single Unit/Combination	37.7/62.3	24.7/75.3
Cargo Body Type		
Cargo/Trailer/Other	73.0/10.2/16.8	74.90/11.9/13.2
Heavy Truck Location of Initial Impact		
Front/Passenger Side/Rear/Driver Side/Other	26.5/25.0/17.0/23.8/7.70	22.5/30.9/4.32/20.0/22.3
Heavy Truck Location of Most Damage		
Front/Passenger Side/Rear/Driver Side/Other	25.5/23.5/15.5/22.5/13	18.9/29.4/4.60/20.5/22.3
Heavy Truck Occupancy		
Continuous	1.12 (0.51)	1.15 (0.50)
Heavy Truck Vehicle Action		
Moving Straight/Turning Left/Turning Right/Passing/Changing Lanes/ Backing/ Slowing/Other/Not Reported/Unknown	55.1/9.82/7.80/1.81/4.14/5.01/5.12/9.89/0.76/0.55	70.5/5.25/11.9/0.33/0.35/2.11/1.52/5.38/2.33/0.33
Vehicle other the a Heavy Truck - Vehicle Age *		
>5 years/>10 years	64.6/35.4	-
Vehicle Other than a Heavy Truck - Location of Initial Impact *		
Front/Passenger Side/Rear/Driver Side	29.1/18.0/13.2/28.4	-
Vehicle Other than a Heavy Truck - Location of Most Damage*		
Front/Passenger Side/Rear/Driver Side	29.0/18.1/12.9/27.9	-
Vehicle Other than a Heavy Truck - Occupancy *		
Single Occupant/Multiple Occupants	74.0/26.0	-
Vehicle Other than a Heavy Truck - Vehicle Action *		
Turning/Slowing/Stopping or Slowing/Other	11.8/6.03/11.0/71.2	-

*Indicates indicator variables established by relating crash level information to the vehicle level. This relate often results in a many to one relationship. Values may not add to 100% due to the possibility of a many to one relationship.

Table 3-1: Summary Statics of Select Variables for Multiple and Single Vehicle Heavy Truck Crashes, 2007-2012 Continued

Variables	Multiple Vehicle Mean (Standard Deviation) or Percentage	Single Vehicle Mean (Standard Deviation) or Percentage
Driver Characteristics		
Heavy Truck Driver's Age		
Continuous	45.5 (13.11)	44.6 (13.3)
Heavy Truck Driver's Gender		
Male/Female/Not Reported/Unknown	90.1/2.75/7.11/0.03	93.5/2.93/2.93/0.04
Heavy Truck Driver Contributing Circumstances		
No Improper Action/Ran Traffic Control Device/Traveling too Fast for Conditions/Crossed Centerline/ Lost Control/Swerved/Operating Recklessly/FTYROW/Distracted/Other/Not Reported/ Unknown	47.9/2.05/2.57/1.72/3.40/1.38/0.68/9.68/0.39/21.7/0.51/8.03	23.3/0.57/9.96/0.53/31.34/3.13/0.90/1.15/1.56/20.2/0.07/7.25
Vehicle Other than a Heavy Truck - Driver's Age*		
<20/<25/<30/40 to 60/>60	12.1/22.8/31.1/27.2/24.5	-
Vehicle Other than a Heavy Truck - Driver's Gender*		
Male/Female/Not Reported/Unknown	51.3/39.8/8.78/0.04	-
Vehicle Other than a Heavy Truck - Driver Contributing Circumstances*		
Traveling too Fast for Conditions/Lost Control/FTYROW	5.83/6.61/9.39	-
Environmental Characteristics		
Weather Conditions		
Clear/Partly Cloudy/Cloudy/Fog or Smoke/Mist/Rain/Sleet/Snow/Severe Wind/Blowing Debris/Other/Not Reported/Unknown	48.5/15.8/11.8/1.40/1.58/4.68/1.77/11.5/0.38/2.06/0.07/0.34/0.24	42.3/12.6/10.3/1.24/1.81/6.92/3.24/10.8/1.67/2.54/0.11/5.80/0.75
Surface Conditions		
Dry/Wet/Ice/Snow/Slush/Sand/Water/Other/Not Reported/Unknown	64.2/11.7/8.20/11.9/1.80/1.31/0.06/0.29/0.31/0.23	55.1/13.0/13.3/7.49/0.93/3.22/0.05/0.44/5.86/0.64
Light Conditions		
Daylight/Dark(road lighted)/Dark(road not lighted)/Other/Not Reported/Unknown	79.8/7.38/8.52/3.97/0.24/0.09	59.14/6.28/23.05/5.87/5.55/0.11

*Indicates indicator variables established by relating crash level information to the vehicle level. This relate often results in a many to one relationship. Values may not add to 100% due to the possibility of a many to one relationship.

Table 3-2: Heavy Truck Crash Distribution, 2007-2012

Vehicle Description	Number of Heavy Trucks in Crashes	Percentage of Heavy Trucks in Crashes
Single-Unit Trucks	**8,735**	**34.9%**
Single-Unit Truck (2-axle/6-tire)	5,732	22.9%
Single Unit Truck (>= 3 axles)	3,003	12.0%
Combination Trucks	**16,268**	**65.1%**
Truck/Trailer	1,669	6.68%
Truck Tractor (bobtail)	270	1.08%
Tractor/Semi-trailer	13,789	55.1%
Tractor/Doubles	264	1.06%
Tractor/Triples	11	0.04%
Other Heavy Truck (cannot classify)	265	1.06%
All Heavy Trucks	**25,003**	**100%**

Table 3-3: Non-Heavy Truck Crash Distribution, 2007-2012

Vehicle Description	Number of Vehicles in Crashes	Percentage of Vehicles in Crashes
Small Passenger Vehicle	**17,851**	**96.94%**
Passenger Car	10,315	56.02%
Four-Tire Light Truck	3,262	17.71%
Van or Mini-Van	1,716	9.32%
SUV	2,558	13.89%
Recreational Vehicle	**129**	**0.70%**
Motor Home	34	0.18%
Motorcycle	82	0.45%
Moped/All-Terrain Vehicle	13	0.07%
Buses	**83**	**0.45%**
School Bus (>15 seats)	30	0.16%
Small School Bus (9-15 seats)	3	0.02%
Other Bus (>15 seats)	41	0.22%
Other Small Bus (9-15 seats)	9	0.05%
Other Vehicle Type	**351**	**1.91%**
Farm Vehicle/Equipment	143	0.78%
Maintenance/Construction Vehicle	28	0.15%
Train	55	0.30%
Not Reported	79	0.43%
Unknown	46	0.25%
All Non-Heavy Trucks	**18,414**	**100%**

3.3 Crash Characteristics

The manner in which a crash occurs, as well as and the number and type of vehicles involved, are significant determinants of the severity outcome of a crash. A distribution of crash severity and vehicle involvement is shown in Figure 3-1. Both multiple and single vehicle crashes show a similar distribution by severity with more severe outcomes being slightly more prevalent in multiple vehicle crashes.

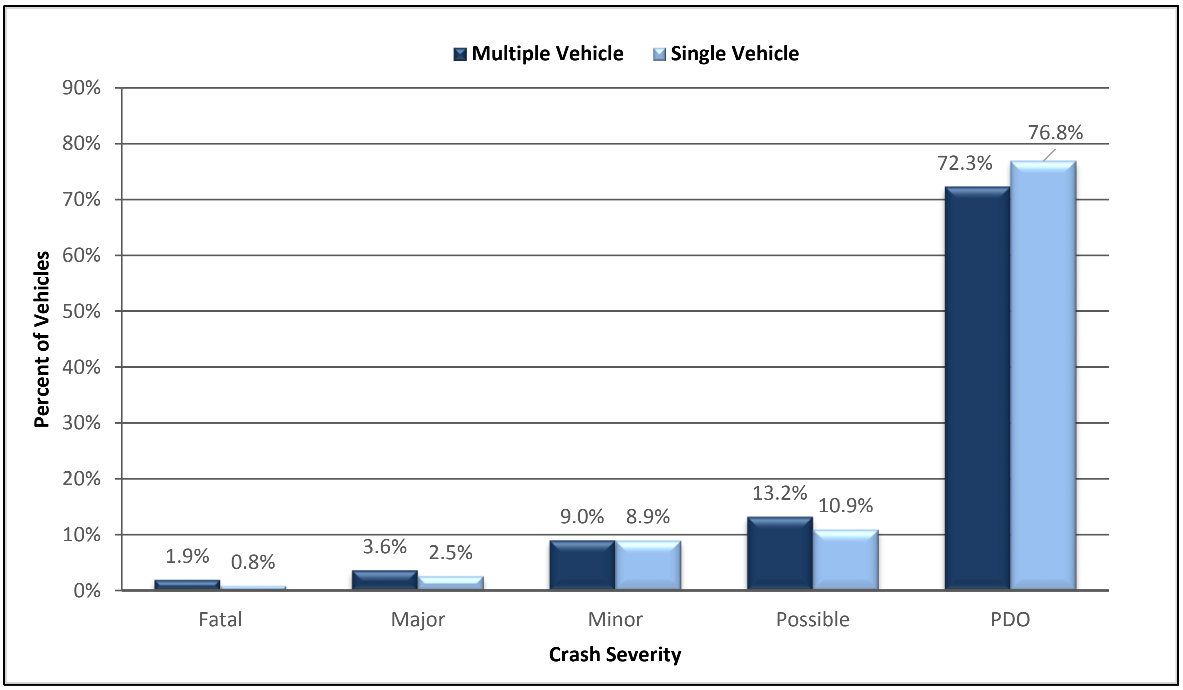


Figure 3-1: Severity Distribution of Single and Multiple Vehicle Crashes, 2007-2012

Though the severity distribution is similar, multiple and single vehicle crashes are quite different with respect to many other crash-specific characteristics. With multiple vehicle crashes there is much greater diversity in the manner in which vehicles collide, as can be seen by comparing Figure 3-2 to Figure 3-3. Sideswipe, rear-end, and broadside crashes tend to be the most common manner of collision for multiple vehicle crashes, while single vehicle crashes are almost explicitly non-collision events.

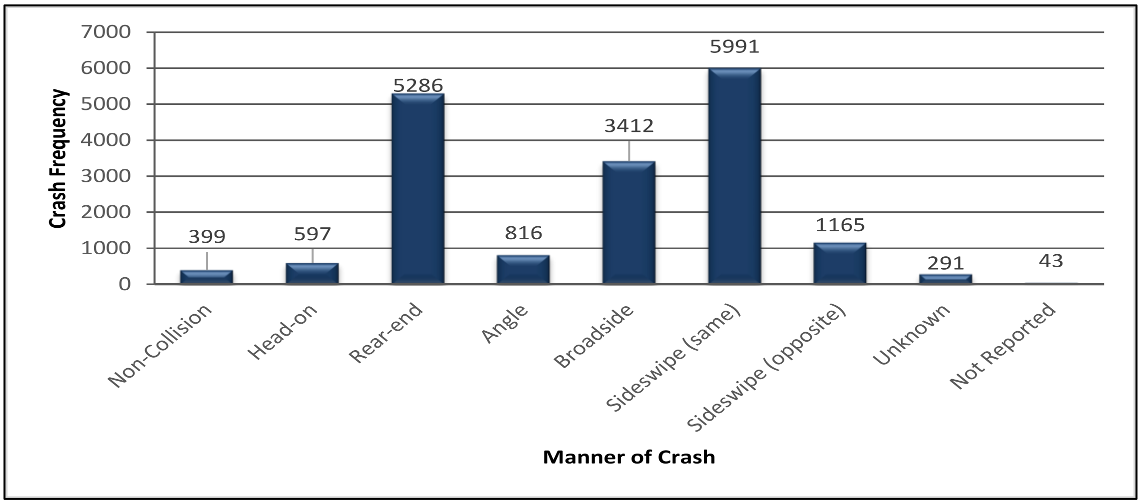


Figure 3-2: Multiple Vehicle Crash: Manner of Crash Frequency Distribution, 2007-2012

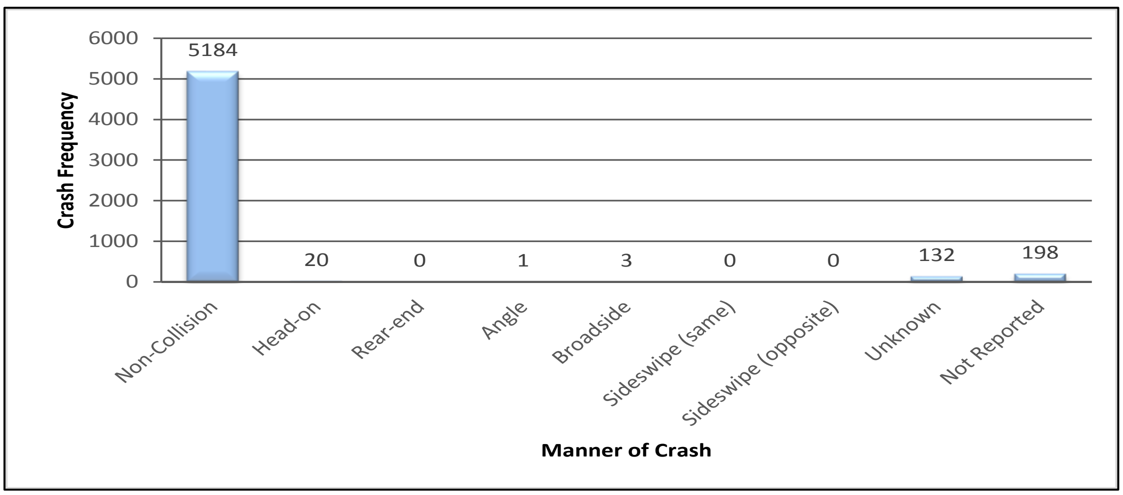


Figure 3-3: Single Vehicle Crash: Manner of Crash Frequency Distribution, 2007-2012

The most harmful event of a heavy truck crash is also likely to be highly related to the severity outcome of the crash. Figure 3-4 and Figure 3-5 show the distribution of the most harmful event reported in multiple and single vehicle crashes respectively. For multiple vehicle collisions the most harmful event is predominately a collision with another vehicle, while for single vehicle collisions the most harmful event is rather variable, with collisions with fixed objects, rollovers, jackknifes, and collisions with animals occurring the most frequently.

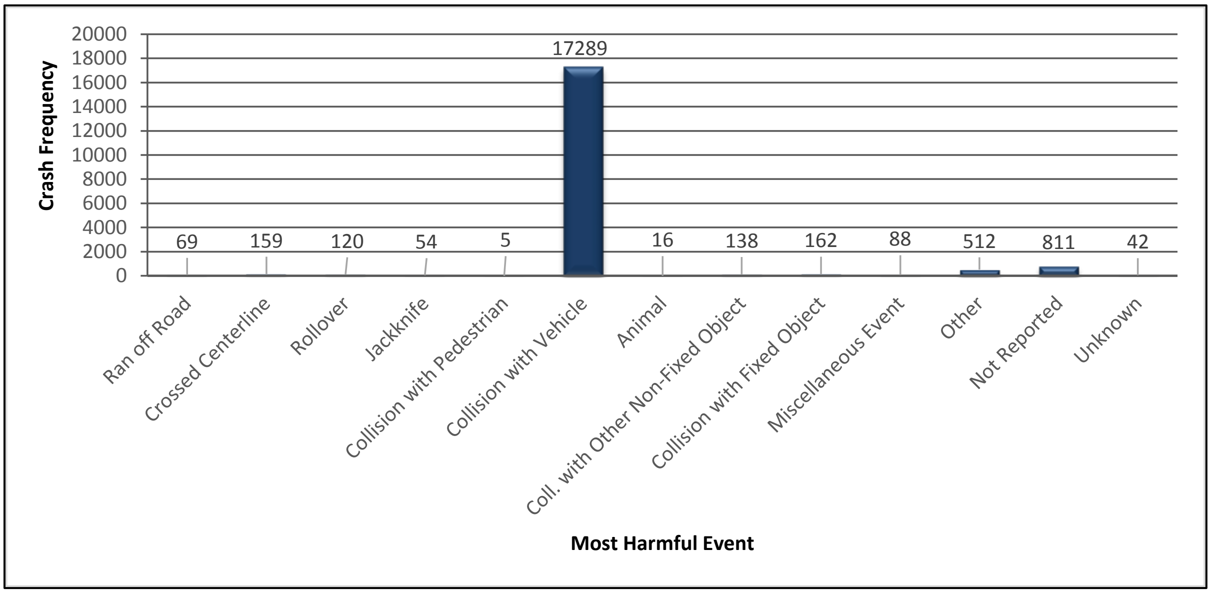


Figure 3-4: Multiple Vehicle Crash: Most Harmful Event Frequency Distribution, 2007-2012

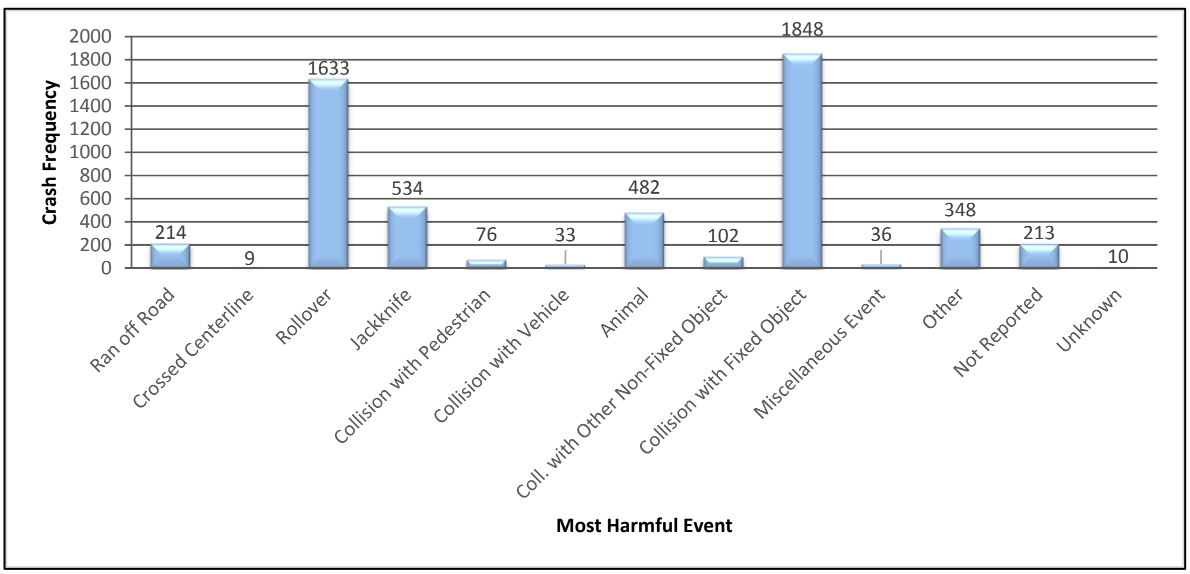


Figure 3-5: Single Vehicle Crash: Most Harmful Event Frequency Distribution, 2007-2012

3.4 Driver Characteristics

As mentioned in the literature review, driver-related factors are commonly cited as the major cause of the crash and as such, a desirable attribute to examine. The data set used for analysis included information on heavy and non-heavy truck drivers' age, gender, condition, crash contributing action, and state of licensure. The gender distribution of heavy truck drivers involved in single and multiple vehicle crashes is almost identical, with male drivers making up over 90 percent of the drivers involved in both crashes. The age distribution of heavy truck drivers involved in a single and multiple vehicle crashes is also similar, with younger drivers appearing to be slightly more involved in single vehicle crashes than multiple vehicle crashes, as can be seen in Figure 3-6. This observation was also verified by a test of proportions ($p < 0.05$, see Appendix A), indicating drivers between the ages of 20 to 34 years old to be over represented in single vehicle crashes. Trends and differences in the age distribution of heavy truck drivers in crashes and the age distribution of all heavy truck drivers in the population were also analyzed. Information on the age of all heavy truck drivers in the state of Iowa was not readily available so as a substitute, the age distribution of drivers getting their commercial driver's license (CDL) renewed between the years 2007-2012 was used to represent the heavy truck driver population. The approximate age distribution of the heavy truck driver population and heavy truck drivers in crashes can be seen in Figure 3-7. For a fair comparison between the CDL data and the crash data, only drivers licensed in Iowa and operating vehicles that require and CDL (all combination trucks) were used for comparison purposes. From the figure one can see that younger drivers appear to be over represented in crashes. This observation was also verified by a test of proportions ($p < 0.01$, see Appendix A), indicating that drivers under the age of 30 were, indeed, overrepresented in crashes.

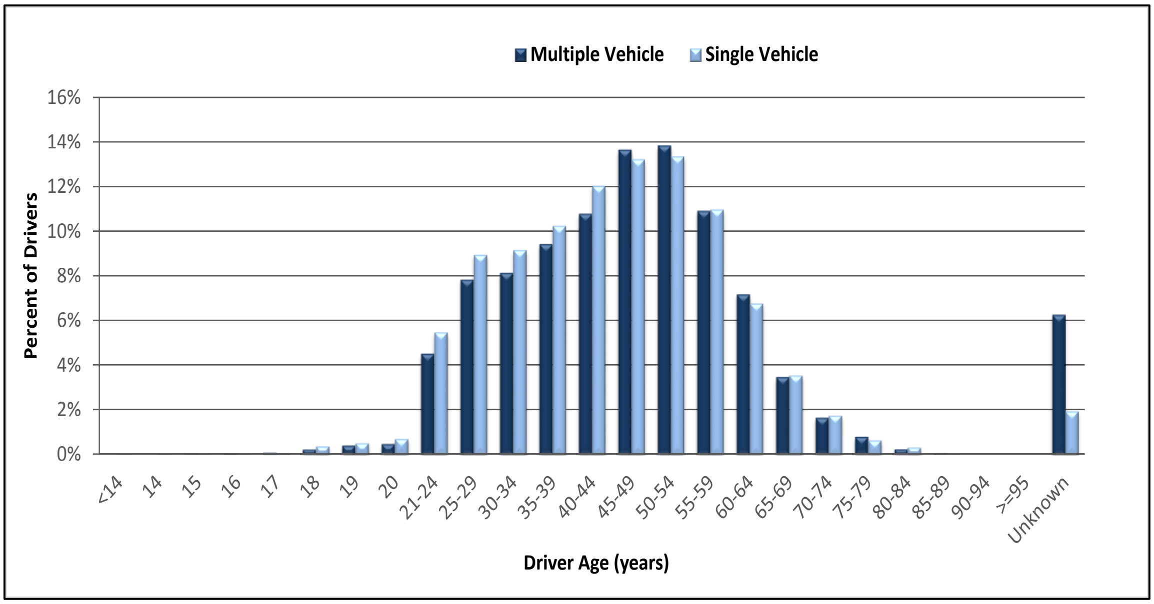


Figure 3-6: Heavy Truck Driver Age Distribution in Multiple and Single Vehicle Crashes, 2007-2012

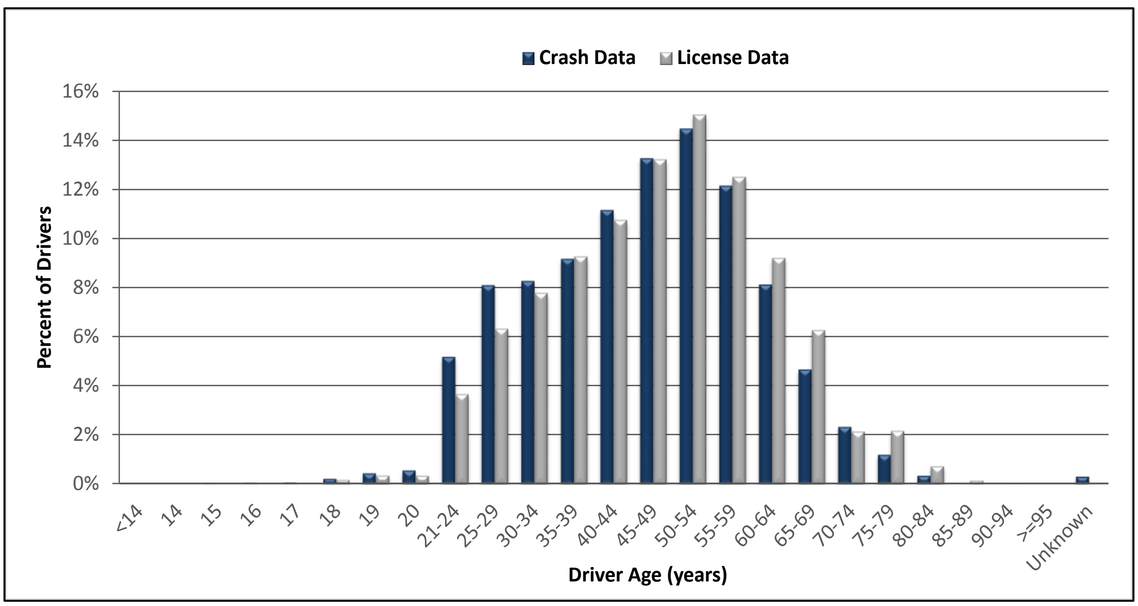


Figure 3-7: Heavy Truck Driver Age Distribution for Drivers in Crashes and Driver's Renewing Their CDL, 2007-2012. (Licensure data obtained through Iowa Motor Vehicle Enforcement)

Both the gender and age distribution of heavy truck and non-heavy truck drivers varies greatly. As can be seen in Figure 3-8, over 90 percent of the heavy truck drivers in crashes are male, while the gender split of the non-heavy truck drivers is close to even. The approximate gender distribution of heavy truck drivers renewing their license and heavy truck drivers in crashes between 2007 and 2012 can be seen in Figure 3-9. Again for a fair comparison between the CDL data and the crash data, only drivers licensed in Iowa and operating vehicles that require and CDL (all combination trucks) were used for comparison purposes. From the figure one can see that the gender distribution of drivers in crashes and drivers renewing their license is similar with males appearing to be slightly over represented in crashes. This observation was also verified by a test of proportions ($p < 0.01$, see Appendix A).

The age distribution of heavy and non-heavy truck drivers is also dissimilar and can be seen Figure 3-10. Non-heavy truck driver's age distribution is widely dispersed with greater representation in both older and younger age groups, when compared to the heavy truck driver age distribution. Heavy truck driver's age distribution is far more concentrated than the non-heavy truck driver's age distribution, with a majority heavy truck drivers being middle-aged. Other driver specific attributes of interest such as alcohol use, drug use, and distraction were reported in such low frequency that it is of little benefit to report such occurrences and attempt to discern a relationship to crash occurrence or crash severity. The temporal and spatial characteristics of heavy truck crashes are discussed next.

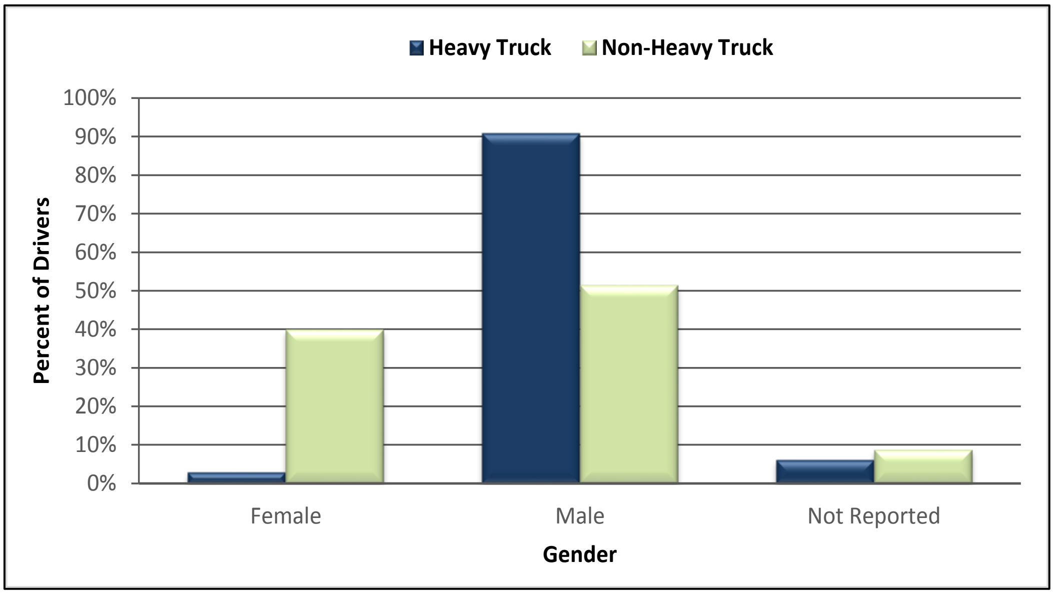


Figure 3-8: Heavy Truck and Non-Heavy Truck Driver Gender Distribution, 2007-2012

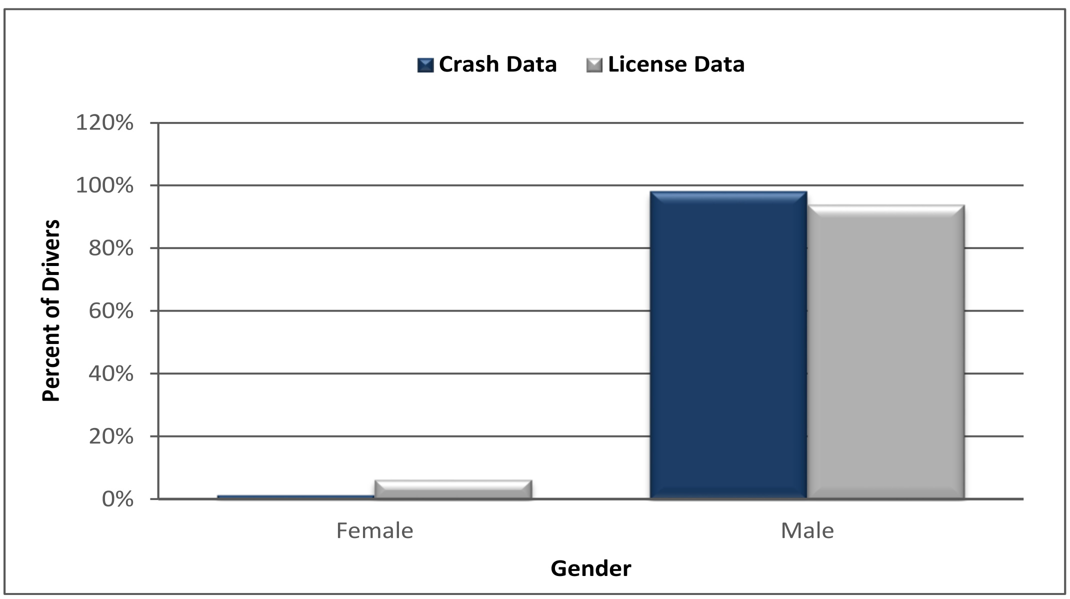


Figure 3-9: Heavy Truck Driver Gender Distribution for Drivers in Crashes and Driver's Renewing their CDL, 2007-2012. (Licensure data obtained through Iowa Motor Vehicle Enforcement)

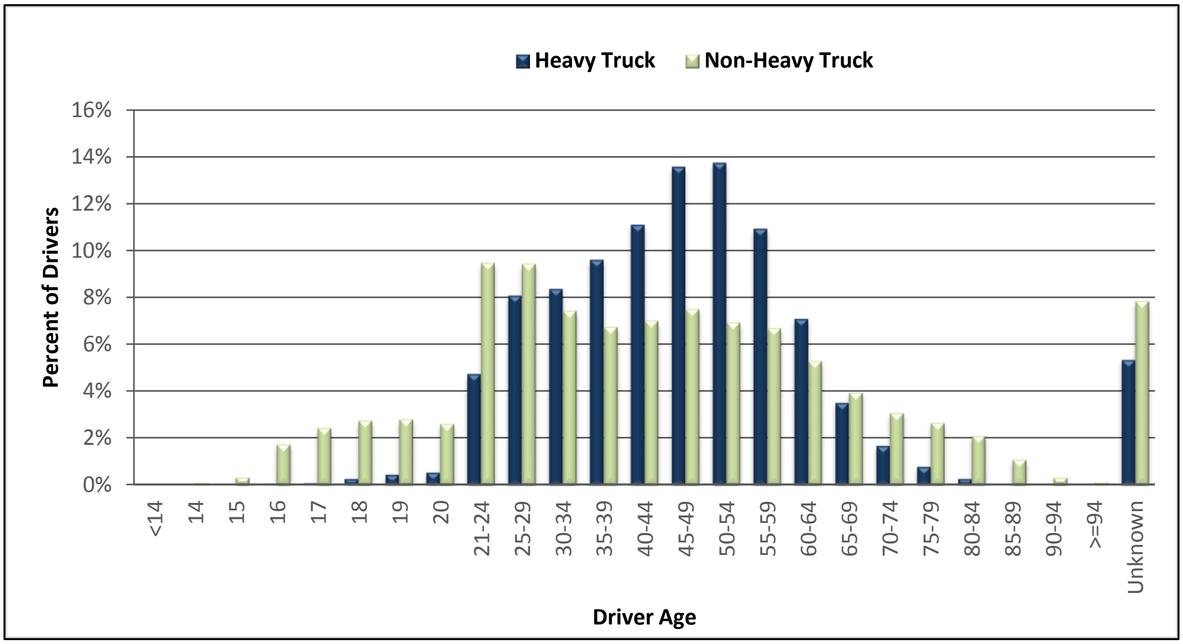
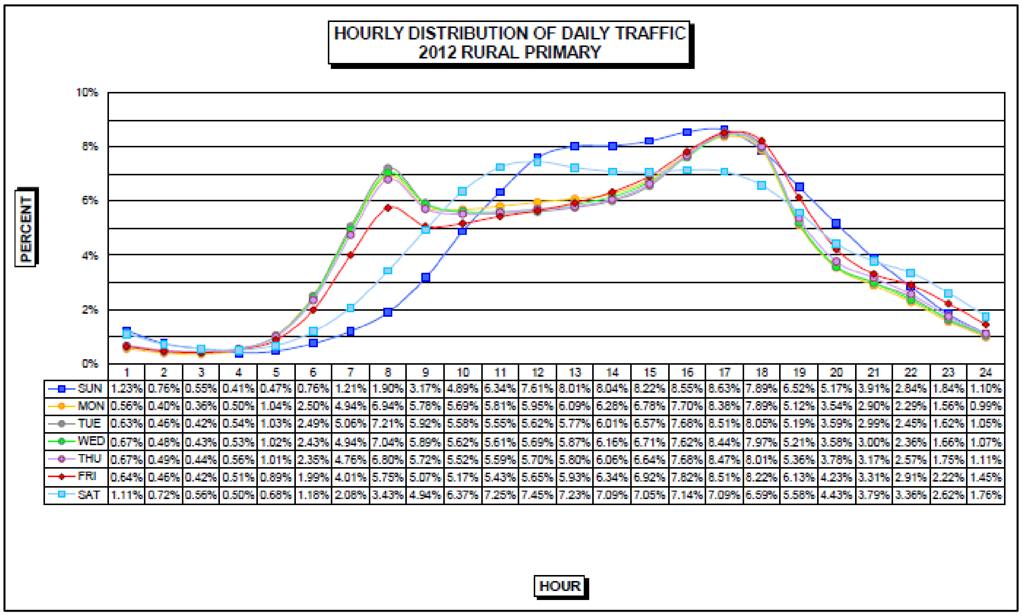


Figure 3-10: Heavy Truck and Non-Heavy Truck Driver Age Distribution, 2007-2012

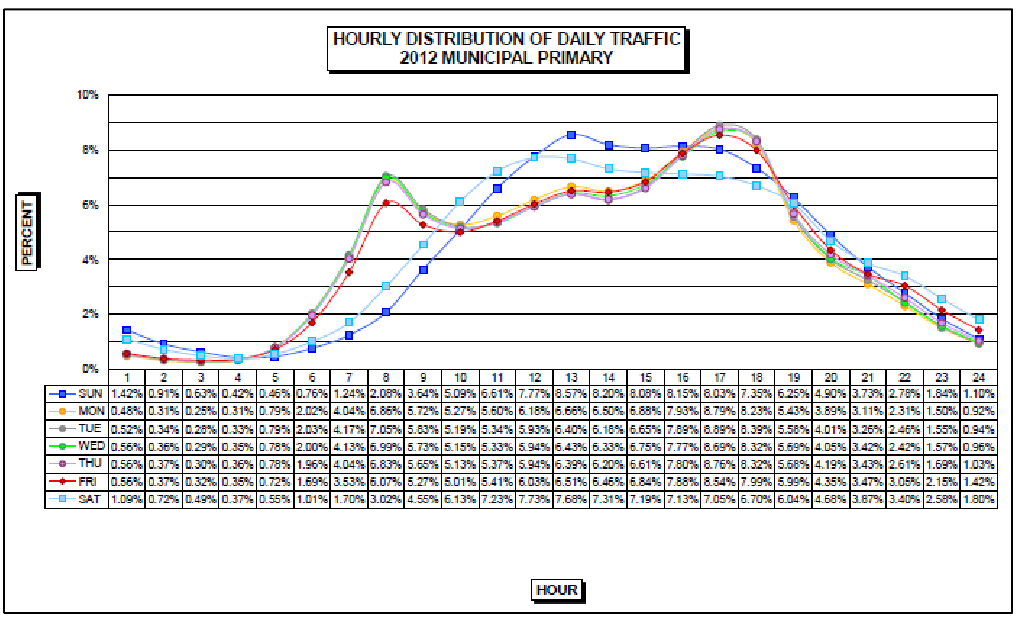
3.5 Time and Location Characteristics

The time and location at which crashes occur is of great importance in the development of appropriate countermeasures. Insight into temporal and spatial trends is also necessary to fully assess safety in a region or associated with a specific demographic group. Traffic on Iowa roadways follows a temporal pattern, with traffic peaking on weekdays during the morning, afternoon, and evening peak hours as can be seen in Figure 3-11 and Figure 3-12. During these peak times of the day the exposure to other vehicles on the roadway is the greatest. As the exposure increases so to should the likelihood of a collision. This trend in exposure needs to be taken into account to when interpreting any trends noticed in the data.



	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
SUN	1.23%	0.76%	0.55%	0.41%	0.47%	0.76%	1.21%	1.90%	3.17%	4.89%	6.34%	7.61%	8.01%	8.04%	8.22%	8.55%	8.63%	7.89%	6.52%	5.17%	3.91%	2.84%	1.84%	1.10%
MON	0.56%	0.40%	0.36%	0.50%	1.04%	2.50%	4.94%	6.94%	5.78%	5.69%	5.81%	5.95%	6.09%	6.28%	6.78%	7.70%	8.38%	7.89%	5.12%	3.54%	2.90%	2.29%	1.56%	0.99%
TUE	0.63%	0.46%	0.42%	0.54%	1.03%	2.49%	5.06%	7.21%	5.92%	5.58%	5.55%	5.62%	5.77%	6.01%	6.57%	7.68%	8.51%	8.05%	5.19%	3.59%	2.99%	2.45%	1.62%	1.05%
WED	0.67%	0.48%	0.43%	0.53%	1.02%	2.43%	4.94%	7.04%	5.89%	5.62%	5.61%	5.69%	5.87%	6.16%	6.71%	7.62%	8.44%	7.97%	5.21%	3.58%	3.00%	2.36%	1.66%	1.07%
THU	0.67%	0.49%	0.44%	0.56%	1.01%	2.35%	4.76%	6.80%	5.72%	5.52%	5.59%	5.70%	5.80%	6.06%	6.64%	7.68%	8.47%	8.01%	5.36%	3.78%	3.17%	2.57%	1.75%	1.11%
FRI	0.64%	0.46%	0.42%	0.51%	0.89%	1.99%	4.01%	5.75%	5.07%	5.17%	5.43%	5.65%	5.93%	6.34%	6.92%	7.82%	8.51%	8.22%	6.13%	4.23%	3.31%	2.91%	2.22%	1.45%
SAT	1.11%	0.72%	0.56%	0.50%	0.68%	1.18%	2.08%	3.43%	4.94%	6.37%	7.25%	7.45%	7.23%	7.09%	7.05%	7.14%	7.09%	6.59%	5.58%	4.43%	3.79%	3.36%	2.62%	1.76%

Figure 3-11: 2012 VMT by Time of Day for Rural Primary Roads in Iowa Source: Iowa Dot Automatic Traffic Recorder Yearly Report for 2012



	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
SUN	1.42%	0.91%	0.63%	0.42%	0.46%	0.76%	1.24%	2.08%	3.64%	5.09%	6.61%	7.77%	8.57%	8.20%	8.08%	8.15%	8.03%	7.35%	6.25%	4.90%	3.73%	2.78%	1.84%	1.10%
MON	0.48%	0.31%	0.25%	0.31%	0.79%	2.02%	4.04%	6.86%	5.72%	5.27%	5.60%	6.18%	6.66%	6.50%	6.88%	7.93%	8.79%	8.23%	5.43%	3.89%	3.11%	2.31%	1.50%	0.92%
TUE	0.52%	0.34%	0.28%	0.33%	0.79%	2.03%	4.17%	7.05%	5.83%	5.19%	5.34%	5.93%	6.40%	6.18%	6.65%	7.89%	8.89%	8.39%	5.58%	4.01%	3.26%	2.46%	1.55%	0.94%
WED	0.56%	0.36%	0.29%	0.35%	0.78%	2.00%	4.13%	6.99%	5.73%	5.15%	5.33%	5.94%	6.43%	6.33%	6.75%	7.77%	8.69%	8.32%	5.69%	4.05%	3.42%	2.42%	1.57%	0.96%
THU	0.56%	0.37%	0.30%	0.36%	0.78%	1.96%	4.04%	6.83%	5.65%	5.13%	5.37%	5.94%	6.39%	6.20%	6.61%	7.80%	8.76%	8.32%	5.68%	4.19%	3.43%	2.61%	1.69%	1.03%
FRI	0.56%	0.37%	0.32%	0.35%	0.72%	1.69%	3.53%	6.07%	5.27%	5.01%	5.41%	6.03%	6.51%	6.46%	6.84%	7.88%	8.54%	7.99%	5.99%	4.35%	3.47%	3.05%	2.15%	1.42%
SAT	1.09%	0.72%	0.49%	0.37%	0.55%	1.01%	1.70%	3.02%	4.55%	6.13%	7.23%	7.73%	7.68%	7.31%	7.19%	7.13%	7.05%	6.70%	6.04%	4.68%	3.87%	3.40%	2.58%	1.80%

Figure 3-12: 2012 VMT by Time of Day for Municipal Primary Roads in Iowa Source: Iowa Dot Automatic Traffic Recorder Yearly Report for 2012

Figure 3-13 and Figure 3-14 show the hourly distribution of multiple and single vehicle heavy truck crashes, respectively. Typically, one would expect the frequency of vehicular crashes to be highest during peak traffic hours, with peaks in the morning, afternoon, and evening as shown in Figure 3-11 and Figure 3-12. Multiple vehicle heavy truck crashes appear to peak throughout the daylight hours between 7am and 5pm, with the frequency of crashes remaining consistent throughout the day, aside from a slight peak in the late afternoon. Single vehicle heavy truck crash frequency is less stable, with the crash frequency peaking throughout the morning peak hours, and varying throughout the remainder of the 24 hour cycle. Also, single vehicle crashes do not display the same level of concentration of crashes around the workday, as is observed for multiple vehicle crashes. Figure 3-13 and Figure 3-14 also display individual heavy truck crash severity outcomes versus the time of day. It can be observed that severe, multiple vehicle crashes, such as fatal and major injury crashes, appear to steadily increase in frequency throughout the day with a prominent peak during afternoon before frequency then declines. Figure 3-14 shows that severe, single vehicle crash occurrence is highly irregular throughout the day, with discernable peaks occurring in the morning, afternoon, and early evening, with the afternoon peak being the most prominent.

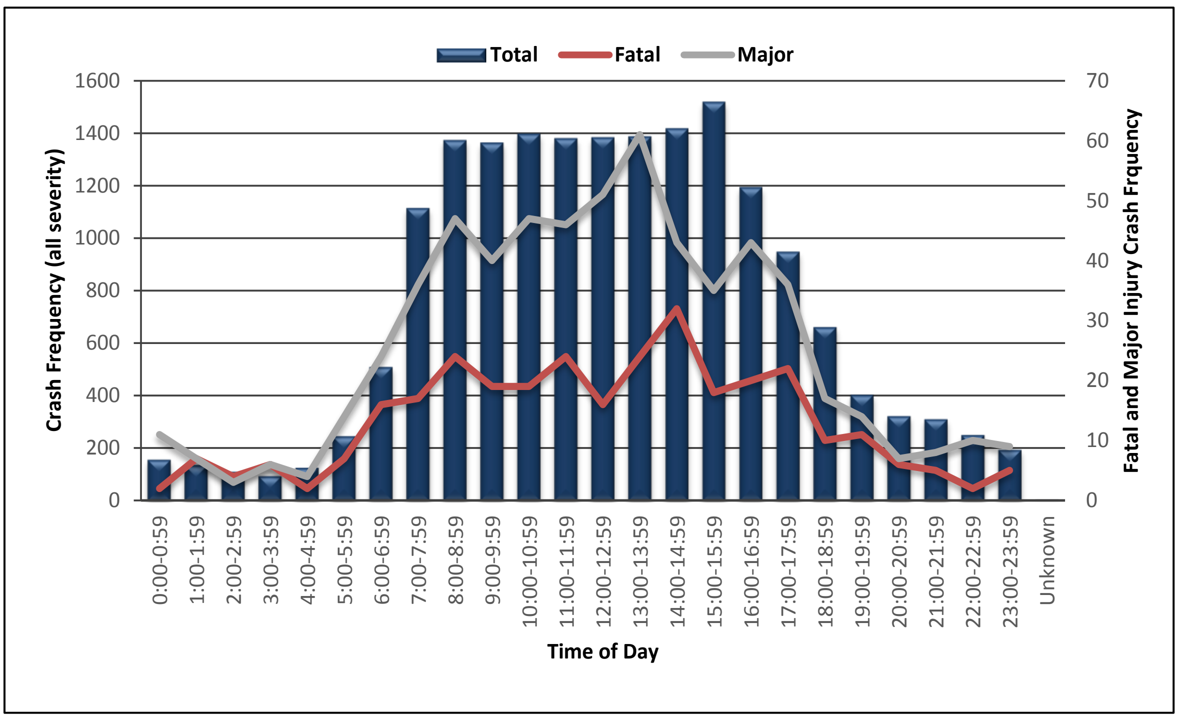


Figure 3-13: Multiple Vehicle Crash Frequency vs. Time of Day, 2007-2012

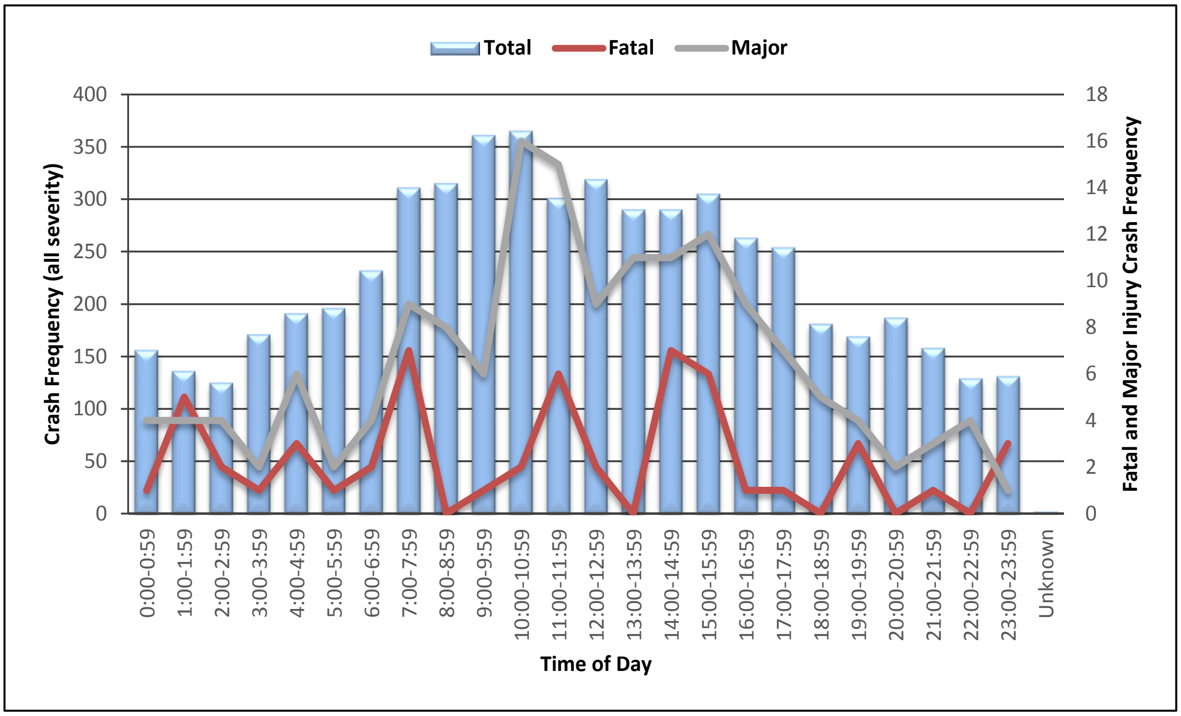


Figure 3-14: Single Vehicle Crash Frequency vs. Time of Day, 2007-2012

Individual days of the week were also taken into consideration. Multiple vehicle and single vehicle crash frequency and their relation to the days of the week can be seen in Figure 3-15 and Figure 3-16, respectively. Figure 3-15 shows that overall, multiple vehicle heavy truck crash frequency tends to be the highest during weekdays, with the crash frequency being fairly stable from Monday to Friday. Similarly, Figure 3-16 shows single vehicle heavy truck crash frequency to be highest during weekdays, but with the frequency of crashes declining as the week progresses from Monday to Friday. From Figure 3-15 it can be seen that severe, multiple vehicle collisions tend to be more frequent toward the beginning of the work week than at the end of the work week. A similar, but much more irregular trend is present for severe, single vehicle collisions, as can be seen in Figure 3-16. To gain further insight into any trends present over the weekend, a test of proportions ($p < 0.01$, see Appendix A), was conducted to see if fatal and major injuries were overrepresented on Saturday or Sunday. The test of proportions concluded that for multiple vehicle collisions, severe crashes were overrepresented on Saturday; however, no significant difference in representation over the weekend was found for single vehicle collisions.

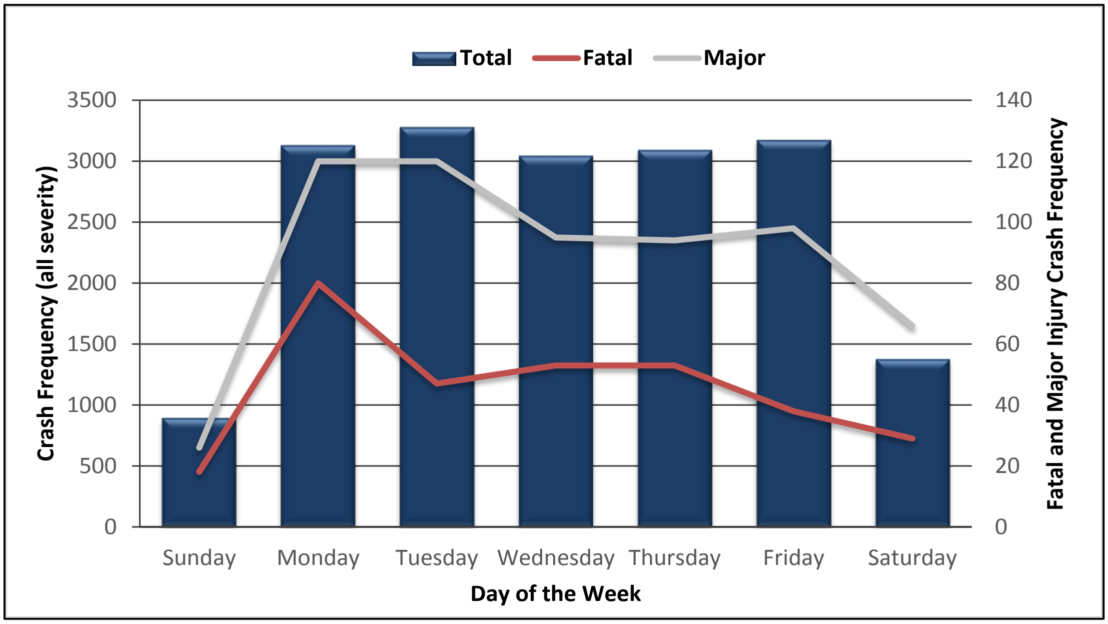


Figure 3-15: Multiple Vehicle Crash Frequency vs. Day of the Week, 2007-2012

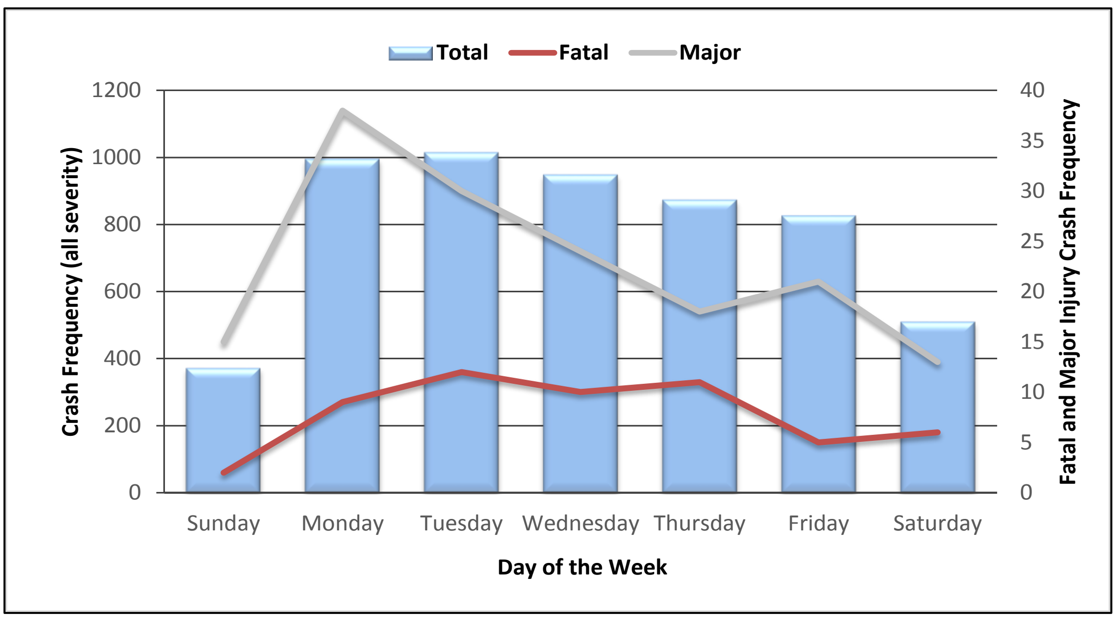


Figure 3-16: Single Vehicle Crash Frequency vs. Day of the Week, 2007-2012

The multiple and single vehicle heavy truck crash distribution by month can be seen in Figure 3-17 and Figure 3-18, respectively. It can be observed that heavy truck crash frequency is highest during the winter months and lowest during the spring, with a slight increase in crash frequency over the summer months. More notable are the differences in the frequency of severe crashes from month to month. Severe, multiple vehicle crashes tend to occur rather irregularly over the year, while severe, single vehicle crash occurrence appears to fluctuate much less from month to month, aside from a prominent peak during the summer months.

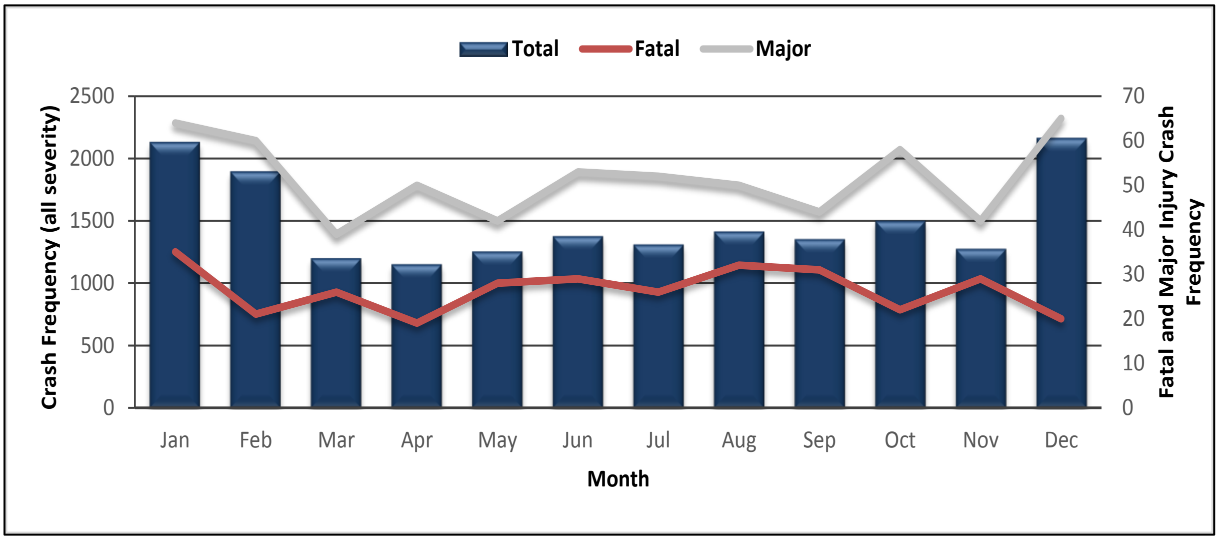


Figure 3-17: Multiple Vehicle Crash Frequency vs. Month, 2007-2012

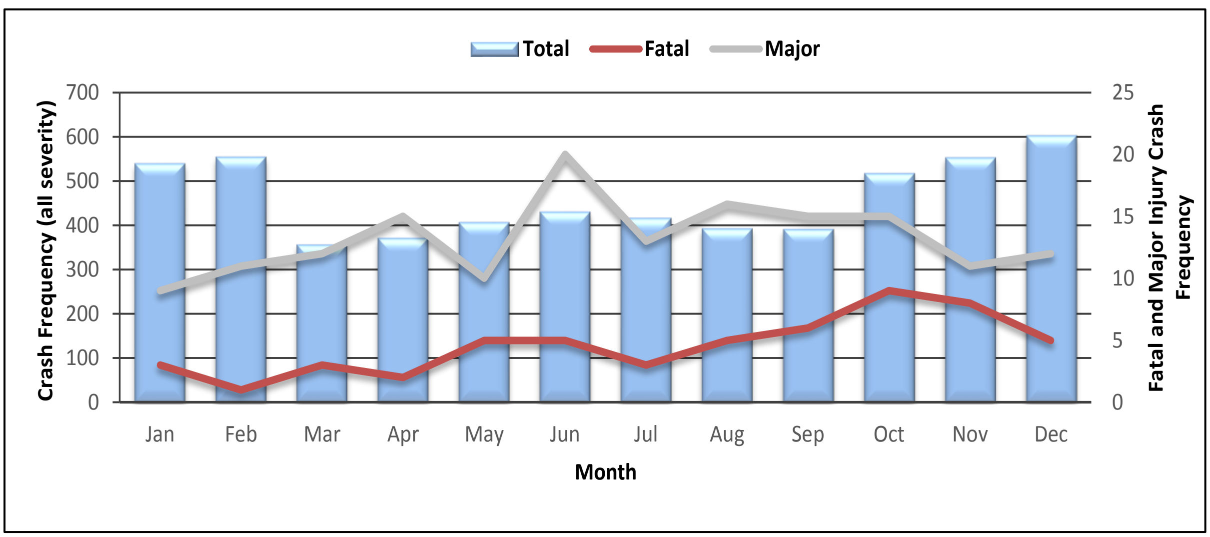


Figure 3-18: Single Vehicle Crash Frequency vs. Month, 2007-2012

The location of a crash is also critical to the complete understanding of heavy truck crash occurrence Figure 3-19 shows the rural and urban crash distribution of multiple and single vehicle heavy truck crashes. It can be observed that single vehicle crashes appear to be predominantly rural events, while multiple vehicle crashes appear to occur most frequently in urban areas. Other factors considered, such as roadway characteristics, are discussed next.

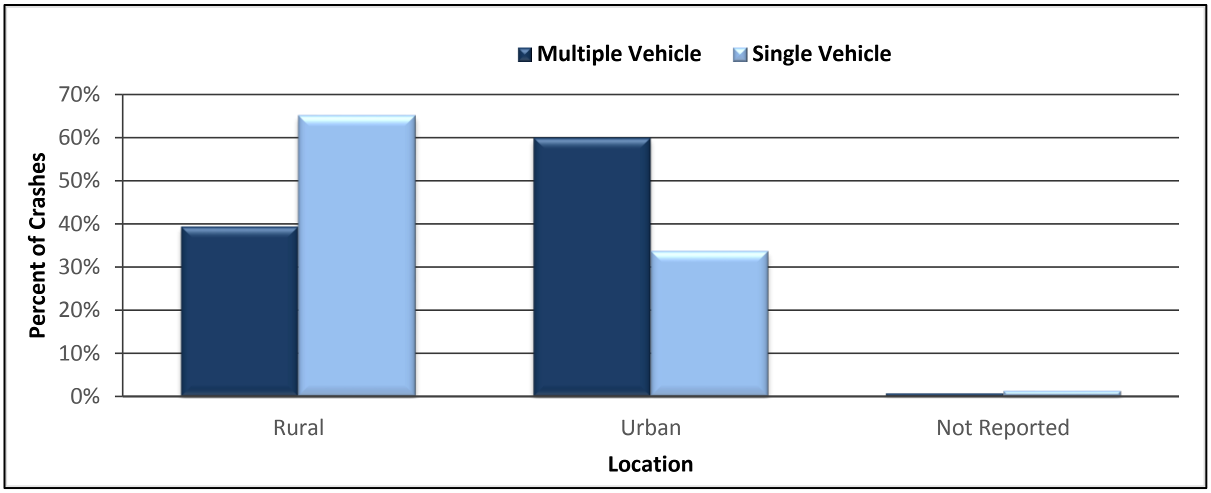


Figure 3-19: Multiple and Single Vehicle Heavy Truck Crash Distribution by Location, 2007-2012

3.6 Roadway and Environmental Characteristics

Information on the type of roadway and characteristics of the roadway where a crash involving a heavy truck occurred were also examined Figure 3-20 and Figure 3-21 show multiple vehicle and single vehicle crash distribution by road classification, respectively. Overall, multiple vehicle crashes occur predominately on municipal roads, interstates, and US routes, with more severe crashes taking place on US routes and interstates. Single vehicle crashes, on the other hand, occur predominately on interstates, secondary roads, and US routes, with the more severe crashes occurring primarily on interstates and secondary roads. The final category of factors considered were environmental characteristics, and they are discussed next.

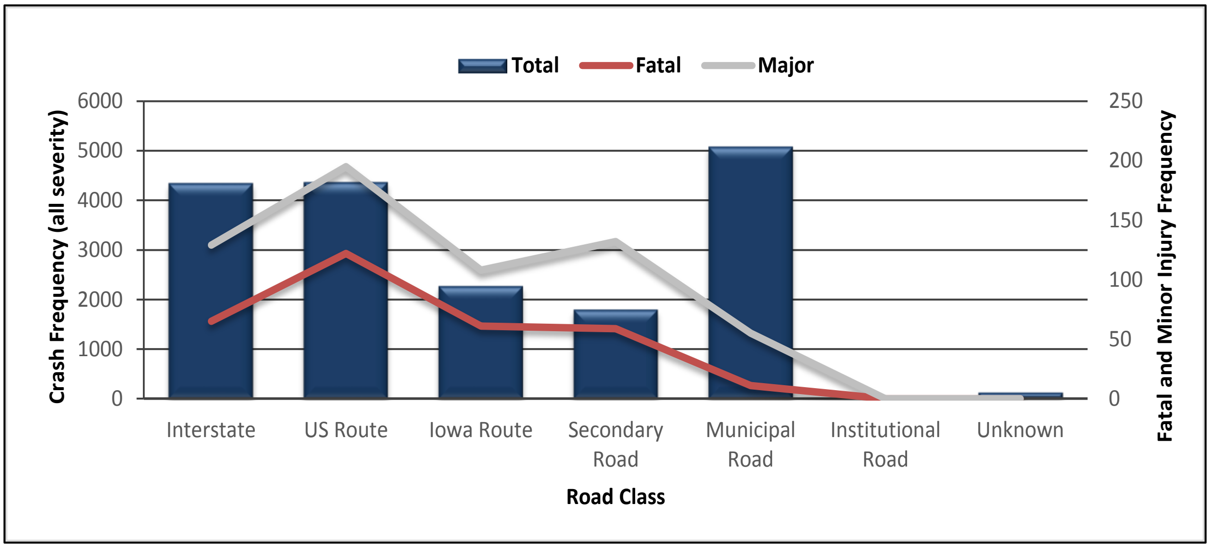


Figure 3-20: Multiple Vehicle Crash Frequency by Road Classification, 2007-2012

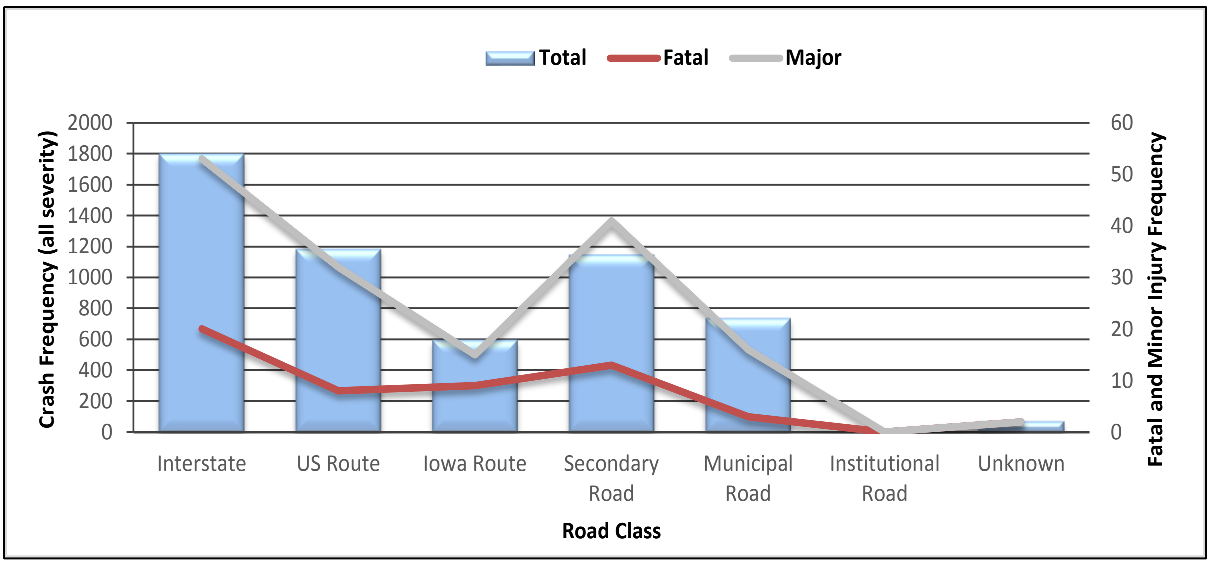


Figure 3-21: Single Vehicle Crash Frequency by Road Classification, 2007-2012

The environmental conditions present at the time of a heavy truck crash are likely to play a role in the frequency and severity of the crash itself. Figure 3-22 shows the crash distribution of multiple and single vehicle heavy truck crashes with respect to the surface conditions present at the time of the crash. From Figure 3-22 it can be seen that a majority of both single and multiple vehicle crashes occur under dry conditions. This observation could be an artifact of the prevalence of dry surface conditions with respect to the other alternative surface conditions reported or related to risk compensating behavior in which drivers drive more aggressively as they perceive dry conditions as safer. Of greater importance is the observation that a higher proportion of single vehicle crashes appear to occur on wet and icy surfaces, while a higher proportion of multiple vehicle crashes occur under snowy and slushy conditions. A test of proportions also supports these observations ($p < 0.01$), see Appendix A).

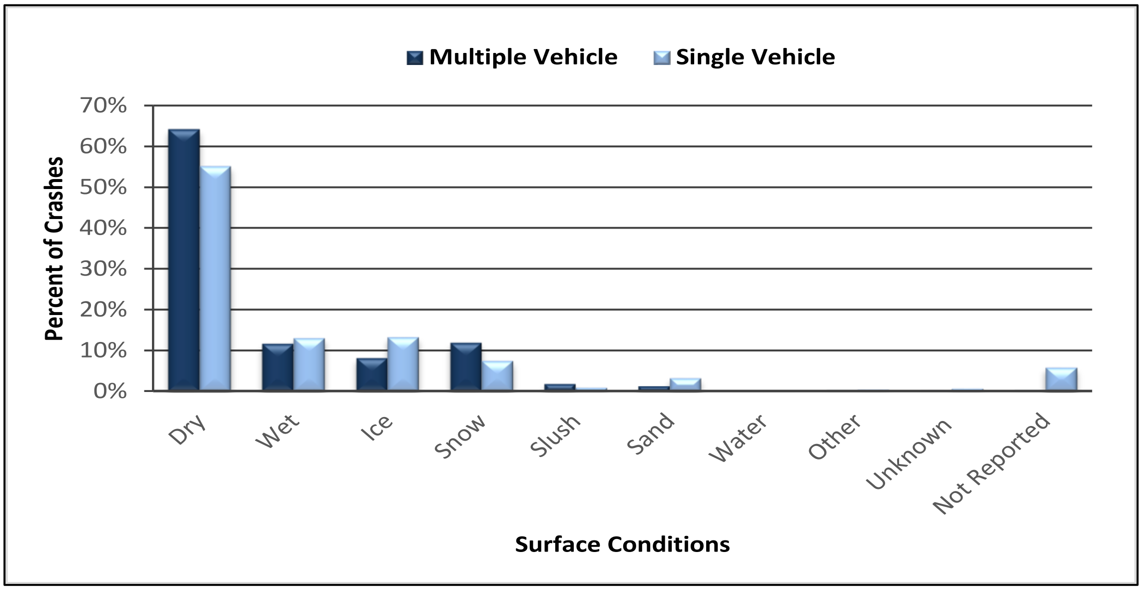


Figure 3-22: Multiple and Single Vehicle Crash Distribution by Surface Condition, 2007-2012

The lighting conditions present at the time of crash occurrence are likely to play a role in the occurrence of a heavy truck crash. Figure 3-23 shows the distribution of multiple and single vehicle heavy truck crashes with respect to the lighting conditions present at the time of the crash. From Figure 3-23, it can be seen that a majority of both multiple and single vehicle crashes occur during daylight lighting conditions with the next highest proportion crashes occurring under dark conditions where the road is not lighted. From the same figure, the disparity of multiple and single vehicle collisions under dark, unlighted, road conditions is rather notable, with a much greater proportion of single vehicle crashes occurring under these conditions as verified by a test of proportions ($p < 0.01$).

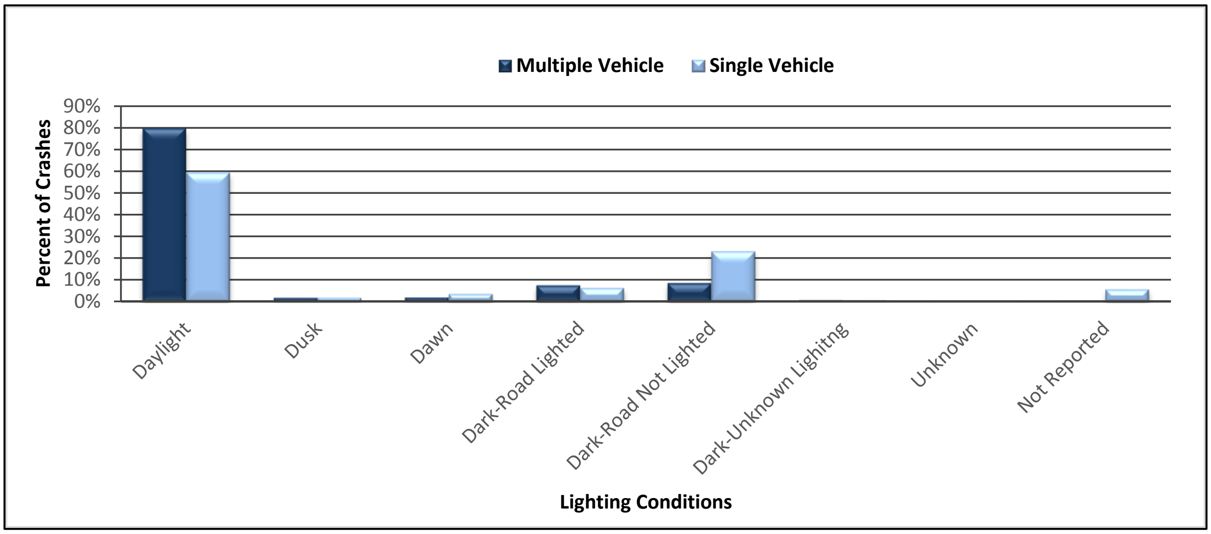


Figure 3-23: Multiple and Single Vehicle Crash Distribution by Lighting Condition, 2007-2012

3.7 Summary

Information on the vehicle, crash, driver, time, location, roadway and environmental characteristics of crashes involving heavy trucks in Iowa from 2007 to 2012 were extracted and analyzed to gain a better understanding of heavy truck crash occurrence in the state of Iowa. From the data, crash frequency and severity trends were observed. A key finding of the descriptive analysis was that there are striking differences between multiple and single vehicle crashes. There are a variety of ways in which vehicles in multiple vehicle crashes collide, while single vehicle crashes are restricted mainly to non-collision events. The most harmful event of a heavy truck crash is also rather different for single and multiple vehicle collisions, with there being a variety of harmful event factors attributable to single vehicle crashes and only a few prevalent in multiple vehicle crashes. Trends in multiple and single vehicle crash time and location factors also followed overall distinct frequency and severity trends. The time of day in which multiple vehicle collisions occur are noticeably concentrated around the daylight hours when people are working. Single vehicle crashes, to the contrary, show a less notable concentration during daylight hours, but with crashes being much more dispersed among other hours of the day. The urban and rural distribution of multiple and single vehicle heavy truck crashes is also very different, with rural crashes being much more prevalent in single vehicle crashes and urban crashes representing a higher proportion of multiple vehicle crashes. Roadway classification was another characteristic in which single and multiple vehicles displayed great disparity. The frequency of multiple vehicle crashes was highest on municipal roads with the greatest concentration of severe crashes occurring on US routes. Single vehicle crashes, on the other hand, appear to occur most frequently on interstates with the highest concentration of severe crashes occurring on interstates and secondary roads. This concludes the discussion on the data set utilized in this thesis. The next chapter will discuss the methodology applied for the statistical analysis.

CHAPTER 4 METHODOLOGY

4.1 Overview of Methodologies Found in Literature

From the literature review, it is clear that the analytical toolbox available for transportation analysis has a vast array of potential tools. The choice of what method or methods to employ is largely dependent on the phenomenon of interest and the data available. However, transportation issues tend to be stochastic nature, which lends well to the use of statistical modeling in transportation analysis.

From the literature review, it was noticed that with-in the context of statistical modeling there is an array of forms a model can take, with the form of the model being largely reliant on the dependent variable of interest and the assumptions necessary for model estimation. Common forms seen in literature include: logistic regression for modeling continuous data such as crash rates; Poisson and negative binomial regression for modeling non-negative integer data such as crash frequency; binary probit or logit, multinomial logit (MNL), and nested logit models for modeling discrete or nominal scale data such as crash severity; and ordered probit models for modeling ordinal discrete data which also includes crash severity.

One of the more frequently used methods of crash investigation in the literature was modeling crash severity using either unordered (multinomial logit or nested logit) or ordered (ordered logit or probit) discrete outcome models. Both ordered and unordered models have their own unique benefits and detriments, and the choice of one method over the other involves taking tradeoffs into consideration. Unordered models, such as MNL models, are susceptible to correlation of unobserved effects from one injury severity level to the other. This correlation violates the MNL model's independence of irrelevant alternatives (IIA) assumption, leading to biased parameter estimates and incorrect inferences (Savolainen et al., 2011). Unordered models also do not account for the ordinal nature of severity data, which decreases the efficiency of parameter estimates. Ordinal models, on the other hand, are not susceptible to IIA violations and do account for the ordering of the data, but are especially susceptible to the underreporting of crashes, resulting in biased and inconsistent parameter estimates (Savolainen et al. 2011). Ordered models also do not possess the flexibility of unordered models because ordered models are unable to capture interior category probabilities (Washington et al. 2011). This means that ordered models place a restriction on variable effect, forbidding the possibility of a variable simultaneously causing probability changes in different directions for other possible outcomes. For example, an ordered model with the outcomes fatal, injury, and no injury may find the effect of air bag deployment to decrease the likelihood of a fatality, which, by nature of the ordered model's constraint of one directional probability shift (see Figure 4-1), would also increase the likelihood of the no injury outcome. However, in reality, it is likely that an airbag deployment would decrease the probability of a fatality and also decrease the probability of no injury (Savolainen et al. 2011). Unordered models are not constrained to ordered model's sometimes unrealistic parameter restrictions and as such can potentially offer a superior fit. Other models suitable for modeling crash severity, not discussed in detail here, include mixed logit models, latent class logit models, or non-parametric models such as classification and regression tree (CART) models, however use of these models is very limited due to complications that arise when interpreting such models' outputs.

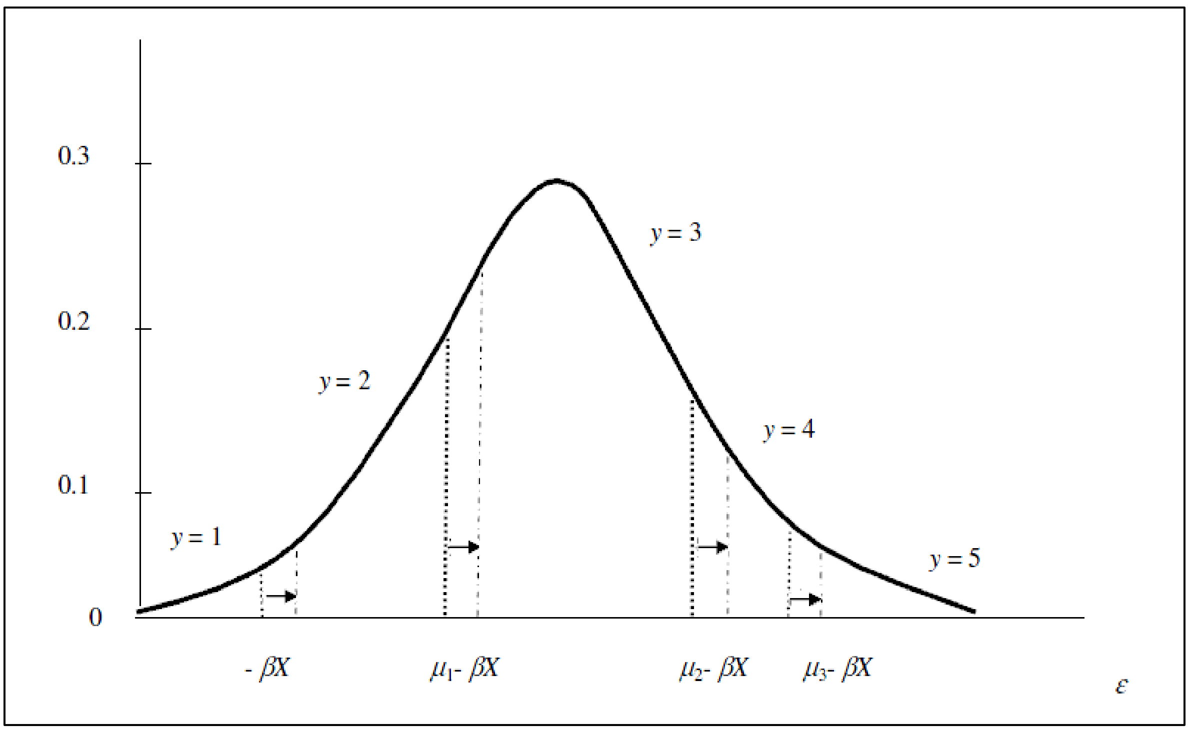


Figure 4-1: Illustration of Ordered Probability Model's One Directional Probability Shift for a Five Outcome Model (Washington et al. 2011)

4.2 Unordered Discrete Outcome Models

From the chapter on data description, the preponderance of differences between single and multiple vehicle crashes is evident. These differences lead to the development of two separate models of heavy truck crash severity, one for single vehicle collisions and one for multiple vehicle collisions. The splitting of the data into two separate models was additionally verified using a transferability likelihood ratio test (see Appendix B).

$$likelihood\ ratio \ = \ -2[LL(\beta_T) - LL(\beta_a) - LL(\beta_b)] \quad (1)$$

where $LL(\beta_T)$ is the log-likelihood at convergence of the model estimated with single and multiple vehicle crash data, $LL(\beta_a)$ is the log-likelihood at convergence of the model using only multiple vehicle crash data, and $LL(\beta_b)$ is the log-likelihood at convergence of the model using only single vehicle crash data. The output of the likelihood ratio test is χ^2 distributed with degrees of freedom equal to the summation of the number of estimated parameters in the single and multiple vehicle models minus the number of estimated parameters in the overall model. A likelihood ratio value greater than the critical χ^2 value supports the use of the separate models over the aggregate model.

Multiple vehicle collisions were modeled using both a three-outcome multinomial logit model and a nested logit model. Single vehicle crashes were modeled using a binary probit model, with the severity outcome of all models being the most severe injury sustained in the crash. The choice to use a binary probit model for single vehicle crashes came as the result of unsuccessful preliminary runs using a three outcome multinomial logit model. Two outcomes (fatal/major injury and minor/possible injury) of the preliminary, single vehicle multinomial logit model, shared many of the same variables with many of the same coefficients. These indifferences in the factors determining the two outcomes suggests a binary model to be a more suitable model for modeling single vehicle crashes. A discussion of multinomial logit, nested logit, and binary probit modeling follows.

4.2.1 Multinomial Logit Model

Multinomial Logit Model Specification

A multinomial logit model for estimating multiple vehicle crash severity outcomes was employed due to the flexibility of the multinomial logit model over the ordered model and its frequent use in similar works previously published. Violations of the IIA property were also tested to ensure the proper functional form was, indeed, utilized. Information on each heavy truck driver and vehicle involved in a multiple vehicle accident in Iowa from 2007-2012 was input into the model. In addition, a restricted set of information on the other vehicle(s) and driver(s) involved in a crash with a heavy truck was joined to each heavy truck crash observation through a relate in ArcMap 10.1 GIS software. For the multiple vehicle crash model, the three injury categories considered were no injury (property damage only), possible or minor injury, and fatal or major injury. The choice of outcome categories was based on observations made during data description and categories used in past studies (Morgan and Mannering, 2011; Gkritza et al., 2010).

For each crash severity outcome the multinomial logit model function takes the form

$$\mathrm{T_{in}} = \mathrm{\beta_i X_{in}} + \mathrm{\varepsilon_{in}} \tag{2}$$

where T_{in} is the function that determines discrete severity outcome i for observation n, β_i is a vector of estimable parameters for severity outcome i, X_{in} is a vector of observable characteristics (driver, crash, vehicle, environment, etc.) that determine the severity of crash n, and ε_{in} is an error term to compensate for possible omitted variables, improper functional form specification, use of proxy variables, and variations in β_i from one observation to the next, not accounted for (Washington et al. 2011).

If ε_{in} are assumed to be extreme value type 1 (Gumbel) distributed the multinomial logit model takes the form

$$P_n(i) = \frac{EXP[\beta_i X_{in}]}{\sum_{\forall I} EXP[\beta_I X_{In}]} \tag{3}$$

where $P_n(i)$ is the probability of crash n resulting in severity outcome i, with I representing the set of all possible crash severity outcomes.

Multinomial Logit Elasticity Estimation

To fully assess the vector of estimated coefficients (β_i), elasticities were computed. Elasticities are a measure of the magnitude of impact a particular variable has on outcome probabilities and can be calculated from the partial derivative of each observation n (n subscripting omitted):

$$E_{X_{ki}}^{P(i)} = \frac{\partial P(i)}{\partial X_{ki}} \times \frac{X_{ki}}{P(i)} \tag{4}$$

where $P(i)$ is the probability of outcome i and $X_{ki}$is the value of variable k for outcome i. Using Eq. (3) and Eq. (4) gives

$$E_{X_{ki}}^{P(i)} = [1 - P(i)]\beta_{ki} X_{ki} \tag{5}$$

where β_{ki} is the estimated coefficient associated with variable X_{ki}. Elasticity values can be interpreted as the percent effect that a 1 percent change in X_{ki} has on the crash severity outcome probability $P(i)$.

Cross-elasticities are a measure of the effect a variable influencing outcome j has on the probability of crash severity outcome i

$$E_{X_{ki}}^{P(i)} = -P(j)\beta_{kj}X_{kj} \tag{6}$$

where *P(j)* is the probability of severity outcome *j* and β_{kj} is the coefficient of variable X_{kj}. Cross- elasticities can be interpreted at the percent effect that a 1 percent change in X_{kj} has on the crash severity outcome probability $P(i)$. It is worth noting that Eq. (6) implies that there is one cross-elasticity for all severity outcomes *i (i≠j).* This property of uniform cross-elasticities is an artifact of the independence of error terms assumed to derive the multinomial logit model (Washington et al., 2011).

Caution needs to be exercised when computing elasticities. Equations (3) to (5) do not apply to indicator variables (variables that take the value of 0 or 1). To gauge the magnitude of the effect of an indicator variable, pseudo elasticities need to be calculated using

$$E_{X_{ij}} = \frac{P_{ij}[given\ X_{ik}=1] - P_{ij}[given\ X_{ik}=0]}{P_{ij}[given\ X_{ik}=0]} \tag{7}$$

where $P_{ij}[given\ X_{ik} = 1]$ is the probability of outcome *i* (direct elasticity) or *j* (cross elasticity) given $X_{ik} = 1$ and $P_{ij}[given\ X_{ik} = 0]$ is the probability of outcome *i* (direct elasticity) or *j* (cross elasticity) given $X_{ik} = 0$. The pseudo elasticity thus represents the percentage change in the probability of a severity outcome when an indicator variable is changed from 0 to 1 (Geedipally et al., 2011).

It should be noted that in case of a variable that is included in more than one utility function, the net effect of the variable can be determined by considering both direct and cross elasticities that are estimated for the variable of interest.

Multinomial Logit Goodness of Fit Measures

The goodness of fit of the multinomial logit model can be found by estimating a ρ^2 statistic using

$$\rho^2 = 1 - \frac{LL(\beta)}{LL(0)} \tag{8}$$

where LL(β) is the log-likelihood at convergence and LL(0) is the log-likelihood with all parameters set to zero. The perfect model would have a ρ^2 statistic equal to one, so the closer the value of ρ^2 is to one, the more variance the model is explaining (Washington et al. 2011). The disadvantage of the ρ^2 statistic is that the value of ρ^2 will always improve with the addition of parameters, regardless of parameter significance. To account for this, an adjusted ρ^2 value can be computed using

$$adjusted\ \ \rho^2 = 1 - \frac{LL(\beta)-k}{LL(0)} \tag{9}$$

where the log -likelihood values are as discussed before and *k* is the number of parameters in the model.

Multinomial Logit Tests of Significance

One final note to consider when interpreting the results of a multinomial logit model is the assumed distribution of error terms. For computational convenience, error term distribution is assumed to be extreme value type 1 (Gumbel) distributed, not normally distributed. This assumption complicates the interpretation of the multinomial logit model's results, but only minimally as is demonstrated in Figure 4-2 by showing the similarity between the assumed logit distribution and the normal distribution.

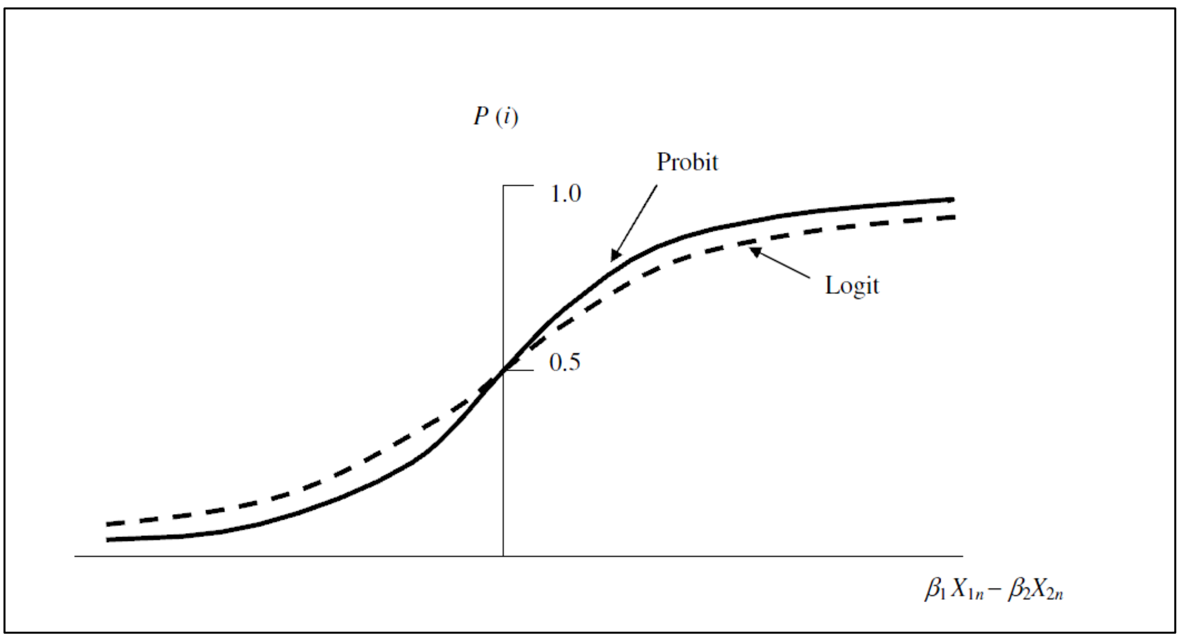


Figure 4-2: Comparison of Binary Logit and Probit Outcome Probabilities (Washington et al. 2001)

For the multinomial logit model the significance of individual parameters is approximated using a one-tailed *t*-test to asses if a parameter is significantly different from zero. The test statistic *t**, which is approximately *t* distributed, is

$$t^* = \frac{\beta - 0}{S.E.(\beta)} \tag{10}$$

where *S.E.(β)* is the standard error of the parameter. Note that because the multinomial logit model is derived using an extreme value distribution, as discussed in the previous paragraph, the use of *t*-statistics is not strictly correct, but a reliable approximation in practice (Washington et al. 2011).

Another more universal and appropriate test for multinomial logit models is the likelihood ratio test. The likelihood ratio test can be used to assess: the significance of individual parameters, a model's overall significance, and the use of separate parameters for the same variable in multiple outcomes (Washington et al. 2011). The likelihood ratio test is

$$\chi^2 = -2[LL(\beta_R) - LL(\beta_U)] \tag{11}$$

where $LL(\beta_R)$ is the log-likelihood at convergence of the restricted model and $LL(\beta_U)$ is the log-likelihood at convergence of the unrestricted model. To test generic attributes, $LL(\beta_R)$ is replaced with the log-likelihood at convergence of the model with generic attributes and $LL(\beta_U)$ is replaced with the log-likelihood at convergence of the model with alternative specific attributes. In either case the likelihood ratio statistic is χ^2 distributed with degrees of freedom equal to the difference in the number of parameters in the restricted/generic model and the unrestricted/alternative specific model.

As mentioned previously, the IIA assumption was also tested to ensure that the assumption holds for the entire choice set and that the proper functional form was, in fact, specified. To check the IIA assumption, Small and Hsiao propose the following test

$$\frac{1}{1 - N_1/(\propto N)}\{-2[LL(\beta_U) - LL(\beta_{UR})]\} \tag{12}$$

Where N is the number of observations in the unrestricted choice set, N_1 is the number of observations in the restricted choice set, and α is a scalar greater than 1 based on the ratio of the covariance matrix of the restricted model and the corresponding elements of the covariance matrix of the unrestricted model. The Small and Hsiao test is χ^2 distributed with degrees of freedom equal to the number of parameters in the restricted model. A test value below the critical χ^2 value confirms the logit model structure cannot be rejected (Ben-Akiva and Lerman, 1985).

4.2.2 Nested Logit Model

Nested Logit Model Specification

Upon completion of the Small and Hsiao IIA test it was discovered that the MNL model structure was not correctly specified. The IIA test revealed that the severity outcomes fatal/major injury and minor/possible injury possibly shared unobserved effects. To resolve this issue a nested logit model, grouping the two previously mentioned severity categories into a conditional nest, was used, see Figure 4-3. The nest structure bears no information on the hierarchy of the decision making process, the nest is simply a method for eradicating IIA violations.

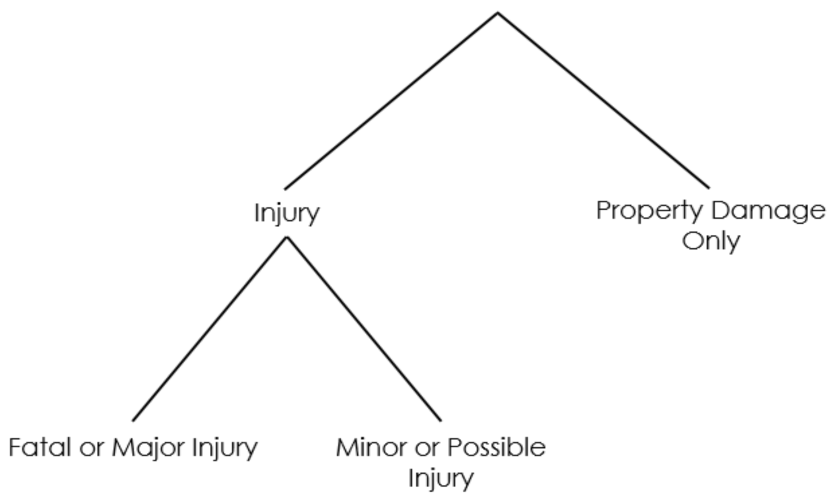


Figure 4-3: Nested Logit Structure for Multiple Vehicle Crash Severity Model

The nested logit model is from the same family of models as the multinomial logit model known as generalized extreme value (GEV) models. The assumption of GEV disturbance terms, allows the IIA problem to be addressed (Washington et al., 2011). Nested logit models group outcomes suspected of sharing unobserved effects into nests. Placing outcomes in nests, allows the shared unobserved effects of the nested outcomes to cancel out, resolving the IIA violation. The nested logit model takes the following model structure

$$P_n(i) = \frac{EXP[\beta_i X_{in} + \varphi_i LS_{in}]}{\sum_{\forall I} EXP[\beta_I X_{In} + \varphi_I LS_{In}]} \tag{13}$$

$$P_n(j|i) = \frac{EXP[\beta_{j|i} X_n]}{\sum_{\forall J} EXP[\beta_{J|i} X_{Jn}]} \tag{14}$$

$$LS_{in} = LN[\sum_{\forall J} EXP(\beta_{J|i} X_{Jn})] \tag{15}$$

$$P_n(j) = P_n(i) \times P_n(j|i) \tag{16}$$

where $P_n(i)$ is the unconditional probability of heavy truck driver *n* having injury outcome *i*, X is a vector of measurable characteristics that determine the probability of injury outcomes, β is a vector of estimable coefficients, and $P_n(j|i)$ is the probability of heavy truck driver *n* having injury severity *j* conditioned on the outcome being in outcome category *i*. For the nest specified in Figure 4-3, *i* is the outcome category injury and $P_n(j|i)$ is the binary logit model of the choice between fatal/major injury and minor/possible injury. Moving on, *J* is the conditional set of outcomes, *I* is the unconditional set of outcomes, LS_{in} is the inclusive value (logsum), and φ_i is an estimable parameter (Washington et al, 2011 and Savolainen & Mannering, 2007).

Nested Logit Tests of Significance

With the nested logit model all previously mentioned calculations for elasticity, pseudo elasticity, goodness of fit measures, and tests of variable significance still apply. One additional test required for the nested model involves the interpretation of the estimated parameter φ_i associated with the inclusive values. φ_i must be greater than zero and less than one in magnitude to be consistent with the nested logit derivation. A *t*-test can be used to test whether φ_i is different from both zero and one. If φ_i is equal to one, then the shared unobserved effects in the nest are not significant and the nested model reduces down to a multinomial logit model. If φ_i is less than zero, then the factors increasing the likelihood of a lower nest, decrease the likelihood of the nest being chosen, which makes no sense and voids the model. Finally, if φ_i is equal to zero, then changes in the outcome probabilities in the nest do not affect the probability of nest selection, suggesting the correct model is separated (Washington et al., 2011).

4.2.3 Binary Probit Model

Binary Probit Model Specification

As mentioned previously, early trials of the multinomial logit, single-vehicle, model revealed that the two outcomes, fatal or major injury and minor or possible injury, shared many of the same variables with the same coefficients. This observation warranted the grouping of all injury outcomes into a single injury outcome, and modeling single vehicle crashes with a binary probit model with the outcomes of injury (fatal or major or minor or possible injury) or no injury (property damage only). Similarly to the multiple vehicle model, information on each heavy truck driver and vehicle involved in a single vehicle accident, not involving a pedestrian, in Iowa from 2007-2012 was input into the model. One key note to make about the binary probit model is that, unlike the multinomial logit model, the binary probit model is not susceptible to an IIA violation and assumes a normal distribution of error terms.

The mathematical form of the binary probit model is

$$P_n(i) = \phi(\beta_0 + \sum_{i=1}^{n} \beta_i X_{in}) \quad (17)$$

$$P_n(j) = 1 - P_n(i) \quad (18)$$

where $P_n(i)$ is the probability of an injury crash, $\boldsymbol{\phi}$ is the standardized cumulative density function, β_o is the intercept, β_i is a vector of estimable parameters, X_{in} is a vector of observable characteristics that determine the severity outcome of crash n, and $P_n(j)$ is the probability of a no injury (property damage only) crash.

Binary Probit Marginal Effects

For the binary probit model, with regard to indicator variables in particular, it is more common to interpret a variables effect using marginal effects instead of elasticities. The marginal effect is the change in absolute probability with respect to a one unit change in the dependent variable and is calculated for continuous variables using

$$Marginal\ Effect\ X_i = \phi(x_i^T \beta) \frac{\delta x_i^t \beta}{\delta x_{ij}} \quad (19)$$

where $\phi(x_i^T\beta)$ is the value of the standard cumulative density function at $x_i^T\beta$ and $\frac{\delta x_i^t \beta}{\delta X_{ij}}$ is the marginal index effect of X_j. For indicator variables the calculation of marginal effects takes the form

$$Marginal\ Effect\ X_j = \phi(x_{1i}^T\beta) - \phi(x_{0i}^T\beta) \tag{20}$$

where $\phi(x_{1i}^T\beta)$ is the value of the probability function with variable X_j equal to 1 and $\phi(x_{0i}^T\beta)$ is value of the probability function with variable X_j equal to zero.

Binary Probit Goodness of Fit Measures

Variable significance and goodness of fit measures for binary models are interpreted and computed in the same way as for the multinomial logit model. Please refer to the earlier discussion on these topics for further information. This concludes the discussion on methodology. The next chapter discusses the results of the modeling methodology just presented.

CHAPTER 5 ESTIMATION RESULTS

5.1 Overview

As discussed in Section 4.2, statistical tests supported the estimation of separate models of single and multiple vehicle heavy truck crash severity. Multiple vehicle crash severity was estimated using both a multinomial logit model and a nested logit model, and single vehicle crash severity was estimated using a binary probit model. This chapter presents an in-depth discussion of the variables found to be significant in both the single and multiple vehicle crash severity models. This chapter also details the effect of many of the variables found to be significant in both the single and multiple vehicle models. Additionally, complimentary findings from the literature review are presented throughout this chapter as a means of fully assessing the estimation results.

5.2 Multiple Vehicle Crash Severity Model

The multinomial logit model was the first model utilized to study heavy truck multiple vehicle crash severity (see Appendix C for results). The specification of a multinomial logit model rests on the assumption of the independence of irrelevant alternatives (IIA), as discussed in Section 4.2 of the methodology chapter. The results of the IIA test indicated that two of the three outcomes of the multinomial logit model shared the same unobserved effects and as such violated the IIA assumption (see Appendix D for calculation). To correct for the IIA violation, a nested logit model was developed where the outcomes fatal or major injury and minor or possible injury were nested to allow their shared unobserved effects to cancel out (refer to Figure 4-3). Table 5-1 presents the estimation results of the nested logit model. A total of 19,465 observations of multiple vehicle heavy truck crashes were used to estimate the model. From the table one can observe the sign and magnitude of each of the 35 variable parameters and two constants included in the model. Parameters with positive signs indicate an increase in the likelihood of a severity outcome, while the opposite effect holds true for negative parameters. The statistical significance of each variable included in the model can also be seen in Table 5-1. A one tailed *t*-test using α=0.05 ($t_{critical}$=1.645) was used to evaluate variable significance. The overall fit of the model is quite good (adjusted ρ^2 of 0.26) given the large amount of variance present in the data set as indicated by the large restricted log likelihood, *LL(0)* equal to -15,695.62. Additional tests of the appropriateness of the nested structure were conducted by verifying the estimated inclusive parameter φ was statistically greater than zero and less than one. This was accomplished using a two tailed *t*-test with α=0.05 ($t_{critical}$=1.96) (see Appendix E for software outputs).

To better interpret the effect of the variables included in the model, elasticities and pseudo elasticities were computed and presented in Table 5-2. As mentioned in Chapter 4, elasticities measure the percent change in the probability of a severity outcome given a one percent change in the value of a variable. Pseudo elasticities, on the other hand, represent the percent change in the probability of a severity outcome given a change in an indicator variable from 0 to 1. All elasticities shown in Table 5-2 are direct elasticities.

5.2.1 Crash Specific Characteristics

For crash specific variables it was found that the manner in which the heavy truck collides has an effect on the severity outcome of the crash. It was desired to include head-on crashes in the study, however the frequency of such crashes did not constitute a sample suitable enough to justify a sole variable for head-on crashes. Instead a variable that included head-on and broadside crashes was created that accounts for both crash types directness of contact and high propensity for damage. Findings show that the variable for head-on or broadside collisions increases the likelihood of an injury occurring by 67 percent. This is anticipated given the high amount of damage typically associated with both crash types. The outcome of a sideswipe crash was found to be less severe than either head-on or broadside crashes, with a sideswipe being 21 percent more likely to result in no injury (PDO). Again this result is not unexpected, but does further validate the results of the model.

Another finding from the multiple vehicle model relates to the number of vehicles involved in a crash. It was found that crashes involving three or more vehicles increased the probability of an injury by 56 percent. This finding seems reasonable. Crashes that involve more vehicles also involve more drivers and passengers. The more people in a crash, the more likely it is for an injury to be sustained by at least one person. This finding is also consistent with past studies (Cheng and Mannering, 1999; Lemp et al., 2010).

Table 5-1: Nested Logit Model Estimation Results for Multiple Vehicle Heavy Truck Crashes

Variable	Injury (Upper Nest)		Fatal or Major Injury		Minor or Possible Injury		PDO	
	Coefficient	*t*-Statistic	Coefficient	*t*-Statistic	Coefficient	*t*-Statistic	Coefficient	*t*-Statistic
Constant								
Minor/Possible Injury Crash	-	-	-	-	2.534	15.08	-	-
Property Damage Only (PDO) Crash	-	-	-	-	-	-	4.163	11.18
Crash Specific Characteristics								
(HDBRD) Head-on or Broadside Crash	0.740	15.08	-	-	-	-	-	-
(SDSWIPE) Sideswipe (same direction) Crash	-	-	-	-	-	-	0.709	13.60
(3PLUS) 3or More Vehicles in a Crash	0.654	10.37	-	-	-	-	-	-
(HTHT) Heavy Truck Crash with Heavy Truck	-	-	0.526	5.51	-	-	-	-
(VAN) Crash Involved a Van	-	-	0.596	5.19	-	-	-	-
(CAR) Crash Involved a Car	-	-	0.307	3.57	-	-	-	-
(SUV) Crash Involved a SUV	-	-	0.443	4.11	-	-	-	-
Time and Location Characteristics								
(LTSUMM) Late Summer (July, August, or September)	-	-	-	-	0.120	2.00	-	-
(FALL) Fall (October, November)	-	-	-	-	-	-	0.159	3.00
(BWEEK) Beginning of the Week (Monday or Tuesday)	-	-	0.282	3.69	-	-	-	-
(EWEEK) End of the Week (Thursday or Friday)	-	-	-	-	-	-	0.087	2.21
(WKND) Weekend (Saturday/Sunday)	-	-	0.289	2.63	-	-	-	-
(AM) Morning (5AM to 8AM)	-	-	0.211	1.95	-	-	-	-
(AFTRN) Afternoon (11AM to 2PM)	-	-	0.232	2.75	-	-	-	-
(PM) Evening Peak (3PM to 6PM)	-	-	-	-	-	-	0.161	3.46
Vehicle Attributes								
(COMB) Cargo Type Combination Truck	-	-	0.823	7.61	0.290	3.38	-	-
(HTFRNT) Heavy Truck Front Initial Impact	-	-	1.362	7.65	0.875	5.07	-	-
(HTSIDE) Heavy Truck Side (driver or passenger side) Initial Impact	-	-	-	-	0.209	2.87	-	-
(PVFRNT) Passenger Vehicle Front Most Damage	0.446	9.88	-	-	-	-	-	-
(PVSIDE) Passenger Vehicle Side Most Damage (driver or passenger side)	-	-	0.236	3.02	-	-	-	-
(PVREAR) Passenger Vehicle Rear Most Damage	-	-	-	-	0.418	4.28	-	-
(PVAGE10) Passenger Vehicle 10+ Years Old	0.302	7.92	-	-	-	-	-	-
(PVMULTIO) Passenger Vehicle had Multiple Occupants	0.140	3.39	-	-	-	-	-	-
Driver Characteristics								
(HTAGE) Heavy Truck Driver Age	-	-	0.006	2.41	-	-	-	-
(PVDRV60) Passenger Vehicle Driver 60+ Years Old	-	-	0.268	3.22	-	-	-	-
(PVFEMALE) Passenger Vehicle Driver is a Female	-	-	-	-	0.505	7.48	-	-
(PVFTYROW) Passenger Vehicle Driver FTYROW	-	-	-	-	-	-	0.248	3.97
Roadway and Environmental Characteristics								
(SPEED55) Speed Limit 55+ (fatal/major)	1.030	25.40	-	-	-	-	-	-
(WINTRD)Winter Road Surface (Ice, Snow, or Slush)	-	-	-	-	0.326	3.37	0.677	8.74
(Precip) Raining or Misting	-	-	-0.443	-2.81			-	-
(Dark) Dark Environment No Road Lighting	-	-	0.481	3.61	0.265	2.52	-	-
Log Likelihood at zero					**-15,695.62**			
Log Likelihood at convergence					**-11,542.43**			
Adjusted ρ^2					**0.26**			
Inclusive Parameter φ					**0.71**			
t-Statistic $\varphi \neq 0$					5.90			
t-Statistic $\varphi \neq 1$					-2.40			

Table 5-2: Nested Logit Model Estimated Elasticities

Variable	Elasticity (%)			
	Injury (Upper Nest)	Fatal/Major Injury	Minor/Possible Injury	PDO
Crash Specific Characteristics				
(HDBRD) Head-on or Broadside Crash	67	-	-	-
(SDSWIPE) Sideswipe (same direction) Crash	-	-	-	21
(3PLUS) 3or More Vehicles in a Crash	56	-	-	-
(HTHT) Heavy Truck Crash with Heavy Truck	-	60	-	-
(VAN) Crash Involved a Van	-	70	-	-
(CAR) Crash Involved a Car	-	32	-	-
(SUV) Crash Involved a SUV	-	49	-	-
Time and Location Characteristics				
(LTSUMM) Late Summer (July, August, or September)	-	-	7	-
(FALL) Fall (October, November)	-	-	-	5
(BWEEK) Beginning of the Week (Monday or Tuesday)	-	29	-	-
(EWEEK) End of the Week (Thursday or Friday)	-	-	-	3
(WKND) Weekend (Saturday/Sunday)	-	30	-	-
(AM) Morning (5AM to 8AM)	-	21	-	-
(AFTRN) Afternoon (11AM to 2PM)	-	23	-	-
(PM) Evening Peak (3PM to 6PM)	-	-	-	5
Vehicle Attributes				
(COMB) Cargo Type Combination Truck	-	111	19	-
(HTFRNT) Heavy Truck Front Initial Impact	-	233	68	-
(HTSIDE) Heavy Truck Side (driver or passenger side) Initial Im	-	-	13	-
(PVFRNT) Passenger Vehicle Front Most Damage	37	-	-	-
(PVSIDE) Passenger Vehicle Side Most Damage	-	24	-	-
(PVREAR) Passenger Vehicle Rear Most Damage	-	-	28	-
(PVAGE10) Passenger Vehicle 10+ Years Old	24	-	-	-
(PVMULTIO) Passenger Vehicle had Multiple Occupants	10	-	-	-
Driver Characteristics				
(HTAGE) Heavy Truck Driver Age*	-	0.26	-	-
(PVDRV60) Passenger Vehicle Driver 60+ Years Old	-	27	-	-
(PVFEMALE) Passenger Vehicle Driver is a Female	-	-	35	-
(PVFTYROW) Passenger Vehicle Driver FTYROW	-	-	-	7
Roadway and Environmental Characteristics				
(SPEED55) Speed Limit 55+ (fatal/major)	112	-	-	-
(WINTRD)Winter Road Surface (Ice, Snow, or Slush)	-	-	21	20
(Precip) Raining or Misting	-	-33	-	-
(Dark) Dark Environment No Road Lighting	-	54	17	-

*Indicates continuous variable, all other variables are indicator varaibles taking the value of either 0 or 1

Other crash specific characteristics considered were related to the vehicle in a collision with the heavy truck. Collisions between a heavy truck and another heavy truck were found to increase the probability of a fatal or major injury by 60 percent. Heavy trucks are heavy, rigid, and large in size, so a crash involving two or more heavy trucks is likely to a very high energy crash and as such very likely to result in a severe injury as predicted by the model. This finding is also consistent with the results found in Lemp et al. (2010).

Collisions between a heavy truck and either a van, passenger car, or SUV were also found to be significant with respect to fatal or major injury crashes. Of all passenger vehicle collisions considered, collisions between vans and heavy trucks displayed the greatest increase in the probability of a severe outcome, with the probability of fatal or major injury increasing 70 percent. This is likely related to vans being higher occupancy vehicles and as such exposing more people to the threat of an injury should a crash occur. It could also be related to the higher probability of a van carrying young children who are more vulnerable and susceptible to sustain severe injuries should a crash occur. Collisions involving a passenger car and a SUV were found to increase the probability of a fatal or major injury by 32 percent and 49 percent, respectively. The lower estimated elasticity for collisions involving a passenger car compared to vans or SUVs may not make intuitive since at first, especially when one considers the smaller size of passenger cars in relation to vans and SUVs, but similar findings have been found in past work (Kockelman & Kweon, 2002). Passenger cars' lower probability of involvement in severe crashes with heavy trucks, with respect to other types of passenger vehicles, could be related to differences in the driving behavior of passenger car drivers with respect to other vehicles, differences in the safety systems present on passenger vehicles, or possibly passenger vehicle's lower probability of rolling over with respect to vans and SUVs in particular.

5.2.2 Time and Location Characteristics

Time and location variables were also found to be significant in the multiple vehicle crash severity model. Seasons, months, days of the week, and times of the day were all taken into consideration. Various days of the week were significantly related to crash severity outcomes. Crashes occurring at the beginning of the week (Monday and Tuesday) were found to increase the probability of a severe crash by 29 percent. This finding is possibly associated to heavy truck drivers being off duty over the weekend and is line with past research by Park and Jovanis (2010), in which it was found that off duty times of more than 46 hours were associated with an increase in crash risk. This finding suggests that educating drivers to be on alert after extended off duty periods, namely after the weekend, could raise driver awareness of this trend and possibly improve driver performance with respect to operations occurring at the beginning of the week.

Similarly, heavy truck crashes occurring over the weekend (Saturday and Sunday) were also found to increase the probability of a fatal or major injury by 30 percent. Descriptive statistics showed that crash frequency declined greatly over the weekend. However, the test of proportions, as discussed in Chapter 3, indicated severe crashes to be over represented over the weekend. A similar finding was also reported in Kockelman and Kweon (2002), but no justification was provided. The increase in the severity of crashes over the weekend could be the result more leisure travel occurring over the weekend. These leisure trips can cover long distances, be to unfamiliar places, and in many cases involve multiple occupants. Any of these factors could increase the probability of the driver making a judgment error that could lead to a crash. The increase in the probability of a severe crash over the weekend could also be related to weekend increases in impaired driving, however this could not be thoroughly investigated due to the low reported frequency of impaired driving observations contained in the data set used.

Other temporal factors considered were the time of day the crash occurred. Both morning (5AM to 8AM) periods and early afternoon (11AM to 2PM) periods of the day were found to increase the probability of a severe crash by 21percent and 23 percent, respectively. The finding of an increase in severe crashes during the early hours of the day are likely related to driver drowsiness as also reported in Barr et al. (2011). The normally accepted status quo is for drowsiness events to occur as the day comes to an end. Research by Barr et al. (2011) however, found heavy truck drivers to also experience fatigue or drowsy events in the early parts of the day, when most people probably assume they are alert. Educating drivers that fatigue events are also frequent during the early hours of the day would greatly increase driver’s awareness of at least one commonly misperceived risk they face daily. The increased probability of severe crashes in the afternoon is more difficult to comprehend. Traffic on Iowa roads has a tendency to peak in the mornings and evenings (refer back to Figure 3-11 and Figure 3-12). During these peaks, exposure to a crash is the highest; however the increased traffic during these peaks may also be linked to more congestion and possibly lower speeds and less severe crashes as a result. During the afternoon, travel speeds are not impeded by congestion, but exposure remains high. This combination of potentially higher speeds and moderately high traffic volume may be why afternoon crashes are more likely to result in a severe outcome.

5.2.3 Vehicle Characteristics

The type of heavy truck involved in a crash was also found to be related to the severity outcome of heavy truck crashes. Multiple vehicle crashes involving a combination truck were found to be more severe than crashes with a single unit truck. Combination trucks, carrying cargo, were found to increase the probability of the occurrence of a fatal or major injury by 111 percent and increase the probability of a minor or possible injury by 19 percent (both values are direct elasticities). A similar result was reported by Khorashadi et al. (2005), in which the authors attributed the increase in severity to the larger size of combination trucks, with respect to single unit trucks, and the fact that combination trucks consist of multiple units.

The impact locations of heavy truck crashes were also significantly tied to the severity outcome of a crash. The results show that crashes, in which the initial impact is made with the front of the heavy truck, are 233 percent more likely to result in a fatal or major injury and 68 percent more likely to result in a minor or possible injury (both values are direct elasticities). The high elasticity of a heavy truck front impacts indicate that great benefits in truck safety can be achieved by improving the front of heavy trucks or minimizing the potential for vehicles to come into contact with the front of a heavy truck. Much research has been done to improve the crash attenuation structures of heavy trucks, particularly with respect to underride, but the findings from this model suggest more can be done to improve the safety of heavy truck frontal impacts.

Areas of most damage on the vehicle (non-heavy truck) colliding with the heavy truck were also determined to be linked to heavy truck crash severity. The probability of an injury being sustained in a crash increased by 37 percent when the front of the vehicle colliding with the heavy truck was the most damaged. Further when the most damage occurred on the side of the vehicle the probability of a fatality of major injury increased by 24 percent. Another notable finding was that when the most damaged area of the non-heavy truck was the rear of the vehicle, the likelihood of a minor or possible injury increased by 28 percent.

These findings indicate safety improvements to non-heavy trucks would also be beneficial to reducing the severity outcome of collisions with heavy trucks. However, the potential impact of improvements to non-heavy trucks, judging by the magnitude of the elasticities, are not as great as the potential impact possibly gained by making improvements to heavy trucks.

Another notable finding from the model, with respect to the characteristics of the vehicles involved, relates to the age of the non-heavy truck. The results of the model indicate that for non-heavy trucks, built more than 10 years before the time of the crash, the probability of an injury being sustained increases by 24 percent. A positive relationship between older vehicles and crash injury was also reported in O'Donnell and Connor (1996). Older vehicles may not possess the same safety features of newer vehicles and as a result occupants may be more susceptible to injury. Also, as vehicles age, there is a higher potential for failures to occur in either the vehicles mechanical or safety systems, which could potentially result in more severe crash outcomes.

The presence of multiple occupants in the non-heavy truck was also found to increase the probability of a severe crash. Estimates from the model indicate that the presents of multiple occupants in a non-heavy truck increases the probability of a severe crash by 10 percent. A similar finding was also reported in Lemp et al. (2010). The increase in severity could be related to drivers being more distracted by the presence of other occupants or the fact that a multiple occupant vehicle in a crash exposes more people to the threat of injury, increasing the odds of a severe injury being sustained.

5.2.4 Driver Characteristics

The only continuous variable found to be significant in the multiple vehicle crash severity model was the variable for the age of the heavy truck driver. The results of the model found increases in driver age to also increase the probability of a fatal or major injury outcome (elasticity of 0.26). This finding is unexpected given the wealth of general transportation research finding younger drivers to be involved in more crashes. However, few studies on heavy trucks have included information on the driver's age in their final results, with none of the studies reviewed for this thesis including the heavy truck driver's age as a continuous variable. Cantor et al. (2010) found younger heavy truck drivers (under 25) to be linked to crash frequency, but no associations were made between age and crash severity. Cheng and Mannering (1999) found younger heavy truck drivers (under 25) to have higher probability of a possible injury outcome. The positive relationship between heavy truck driver age and fatal or major injury crashes found in this study is more likely related to the physiological differences of younger and older drivers and not the frequency of crash involvement. Younger drivers, in comparison to older drivers, are more resilient in crashes and as such, less likely to sustain a major or fatal injury. Further, crash severity studies, not specific to heavy trucks, have also found older age to be linked to more severe crashes (Abdel-Aty, 2003, O'Donnell & Connor, 1996; Kockelman & Kweon, 2002).

Related to what was just discussed, non-heavy truck drivers, over 60 years old, were found to have higher probability (by 27 percent) of a fatal or major injury when in a crash with a heavy truck. This finding, just as with heavy truck drivers, is linked to older driver's physicality and their susceptibility to more severe injuries and is consistent with many other past studies on crash severity, as discussed above.

The results of the model also found gender to be significant in the determination of crash severity. Female drivers of non-heavy trucks were found to increase the probability of a minor or possible injury 35 percent. This is likely related to the driving behavior of women, in comparison to men. One could speculate women have a tendency to be less aggressive drivers which makes women less prone to severe injuries.

5.2.5 Roadway and Environmental Characteristics

Crashes occurring on roadways with a speed limit over 55 mph were found more likely (by 112 percent) to result in an injury. Heavy truck's greater size, weight, and associated performance limitations, specifically stopping distance, restrict a trucks ability to avoid crashes at higher speeds. This finding along with supplemental findings from Cheng and Mannering (1999) suggest that higher speed limits have a potentially great impact on the severity of heavy truck crashes.

Environmental factors were also found to impact heavy truck crash severity. Winter roadway conditions such as snow, slush, or ice were found to increase the probability of a minor or possible injury by 21 percent and no injury by 20 percent (both values are direct elasticities). A 33 percent decrease in the probability of a fatal or major injury crash was estimated under rainfall events. These findings of a positive relationship between adverse roadway conditions and lower severity outcomes (or negative relationship to severe outcomes) are consistent with past research findings (Lemp et al., 2010; Bham et al., 2012).

The decline in severity under adverse roadway and weather conditions is attributable to drivers being more cautious and attentive under such conditions.

Dark roadway environments with no lighting were found to increase the probability of a fatal or major injury crash by 54 percent and a minor injury crash 17 percent. This finding is similar to that found in the literature (Lemp, 2010; Abdel-Aty, 2003) and is attributable to higher speed variations present on roadways under such conditions as well as the impact of dark roadway environments on a driver's ability to make judgments and respond properly to potential hazards.

This concludes the discussion of the multiple vehicle nested logit model results. The remainder of this chapter discusses the results of the single vehicle binary probit model of crash severity.

5.3 Single Vehicle Crash Severity Model

Single vehicle heavy truck crash severity was first estimated using a multinomial logit model. Initial model outputs of the multinomial logit model indicated that all injury categories (fatal, major, minor, and possible injuries) should be grouped, and that a two-outcome binary model was more suitable for modeling single vehicle heavy truck crash severity. A total of 5,534 observations of single vehicle heavy truck crashes were included in the original data set; 72 of these observations were observations of single vehicle crashes involving a collision between a heavy truck and a pedestrian. These types of crashes, though severe, were not of primary interest in this study and as such, the pedestrian crashes were removed from the data set used for model estimation (leaving 5,462 observations).

Table 5-3 presents the estimation results of the single vehicle binary probit model of crash severity. Due to the binary model structure and the lower number of observations of single vehicle crashes relative to multiple vehicle crashes, fewer significant variable were found in this model. Table 5-3 shows the sign and magnitude of each of the 13 variable parameters and the constant included in the model. Positive coefficients indicate an increase in the likelihood of a crash with an injury sustained, while negative signs indicate the opposite effect. The statistical significance of each parameter included in the model was evaluated using a one tailed *t*-test and α=0.05 ($t_{critical}$ = 1.645). The overall fit of the single vehicle model (adjusted ρ^2 of 0.16) is not as good as the fit of the multiple vehicle model.

Table 5-3: Single Vehicle Binary Probit Model Results

Variable	Coefficient	*t*- Statistic	Marginal Effect
Constant			
Injury Crash	-1.236	-12.70	-
Crash Specific Characteristics			
(X4) Heavy Truck Ran off the Road	0.130	2.85	0.04
(X6) Most Harmful Event was Rollover	0.802	16.45	0.25
(ANML) Most Harmful Event was Hitting an Animal	-0.769	-5.62	-0.16
Time and Location Characteristics			
(X88) Summer (June, July, or August)	0.110	2.10	0.03
(X96) End of Week (Thursday or Friday)	-0.100	-2.15	-0.03
Vehicle Attributes			
(X45) Vehicle was a Single Unit Truck	0.240	4.94	0.07
(X40) Front of Vehicle Most Damaged	0.264	3.87	0.08
(SDDMG) Side of Vehicle Most Damaged	-0.175	-3.29	-0.05
Driver Characteristics			
(X9) Heavy Truck Driver's Age*	0.004	2.83	0.0013
(X19) Heavy Truck Driver Lost Control	0.411	4.48	0.13
(X20) Heavy Truck Driver Traveling Too Fast	0.274	3.50	0.08
Roadway and Environmental Characteristics			
(X29) Speed Limit is less than 35mph	-0.472	-7.33	-0.12
(X86) Winter Surface Conditions	-0.423	-7.17	-0.11
Log Likelihood at zero		**-2726.43**	
Log Likelihood at convergence		**-2290.03**	
Adjusted ρ^2		**0.16**	

*Indicates continuous variable, all other variables are indicator variables taking the value of either 0 or 1

The single vehicle model's inferior fit, in comparison to the multiple vehicle model, is likely due to the fewer number of variables that were introduced in the model (for example information on the non-heavy truck driver and vehicle), and found to be significant in the multiple vehicle model. Additionally, some of the most explanatory variables included in the multiple vehicle model such as the manner of collision, were not applicable to the single vehicle model, leaving fewer variables available to explain the variance of the data. Further, the data set used does not contain a lot of driver specific variables and such variables are likely the cause of single vehicle crashes (see Appendix F for software outputs).

To better interpret the results of the single vehicle binary probit model, it is common practice to estimate marginal effects for each variable included in the model instead of elasticities. Marginal effects represent the absolute change in probability for a unit change in an independent variable. Please refer to Table 5-3 for the results of this estimation.

5.3.1 Crash Specific Characteristics

The estimation results suggest that when a heavy truck runs off the road there is a 0.04 higher probability of an injury. This finding is also consistent with findings in Cheng and Mannering (1999).

Findings from the model also suggested that the occurrence of a rollover increases the probability of an injury by 0.25. This is a large increase in the likelihood of an injury and suggests large impacts to truck safety can be made through measures designed to prevent rollovers or reduce the severity of a crash should a rollover occur. Potential countermeasures include mandating all truck to have electronic stability control or, as suggested in Perrin et al. (2007) equipping trucks with side airbags. Additionally, training or education on controlling a heavy truck could raise awareness of the severity of such events and, as a result, reduce the occurrence of heavy truck rollovers.

Collisions with animals were the only crash related factor found to decrease the likelihood of an injury. The probability of an injury in collisions with animals was found to be decrease by 0.16. This finding indicates that animal-heavy truck collisions are not of much consequence (as collisions of passenger vehicles with a animals) and heavy trucks are presently equipped well enough to resist injury to the driver should a heavy truck come into contact with an animal on the road.

5.3.2 Time and Location Characteristics

Temporal factors were also found to be significant in the determination of single vehicle crash severity. Crashes during the summer (June, July, or August) were found to increase the probability of an injury by 0.03. This finding could be in relation to driver drowsiness and fatigue that come as a result of the higher temperatures present during the summer months.

Crashes toward the end of the work week (Thursday and Friday) were found to decrease the likelihood of an injury by 0.03. This could be related to drivers getting into a driving rhythm as the week progresses or to drivers, with regular commutes, becoming more accustomed to their route as the week progresses. A similar finding was also found in the multiple vehicle crash severity model.

5.3.3 Vehicle Characteristics

One finding from the single vehicle crash severity model, contrary to the multiple vehicle model, was that crashes involving single unit trucks, not combination trucks, were more likely to result in an injury. Single unit trucks, in single vehicle crashes, were found to increase the probability of an injury by 0.07. To understand the differences between the two models, it is best to again consider the difference between multiple and single vehicle crashes. Combination trucks are bigger and heavier than single unit trucks and as a result can cause more damage when in collisions with other vehicles. In multiple vehicle crashes, this means that combination trucks damage the other vehicle in the collision more than single unit trucks would, which likely causes more severe injuries to the occupants of the other vehicles. In single vehicle crashes, it is only the heavy truck that sustains damage and only the occupants of the heavy truck at risk for injury. Single unit truck’s smaller size might make them more susceptible to damage and injury. Single unit truck’s positive relationship to crash severity in single vehicle crashes may also be linked to the characteristics of drivers of single unit trucks versus those of drivers of combination trucks. Combination trucks require a commercial driver’s license (CDL) to operate, while single unit trucks do not necessarily require a CDL. As such, it can be assumed that drivers of combination trucks are better trained in vehicle control than single unit drivers and this could probably result in a higher occurrence of an injury.

Other vehicular factors found significant in single vehicle collisions are in relation to the part of the heavy truck that experienced the most damage. Crashes in which the heavy truck's front experienced the most damage increased the probability of an injury by 0.08 while damage to either the driver or passenger side of the truck decreased the probability of an injury by 0.05. These findings reinforce what was mentioned earlier in the discussion of the multiple vehicle model estimation results. To improve heavy truck safety, attention to improving the front of the heavy truck with respect to protecting the occupants of the other vehicle, as discussed before, and with respect to protecting the occupants of the heavy truck, as found in this model, would greatly decrease the severity of all heavy truck crashes.

5.3.4 Driver Characteristics

The age of the heavy truck driver was the only continuous variable found significant in the single vehicle crash severity model. It was found that a unit increase in the age of the driver increased the probability of an injury by 0.0013. As was discussed earlier this relationship has been found in many past studies, (Abdel-Aty, 2003; O'Donnel & Connor, 1996; Kockelman & Kweon, 2002) and is likely linked to the physiological differences of older and younger people, with older drivers being more susceptible to injury.

Other interesting driver attributes found significant in single vehicle crashes relate to the drivers' operation of the heavy truck. It was found that crashes in which the driver was reported as traveling too fast for conditions or speeding, the probability of an injury increased by 0.09. This finding is complimentary to findings in Cheng and Mannering (1999) and Lemp et al. (2010), and suggests that education and enforcement measures to prevent speeding could be effective at improving heavy truck safety.

5.3.5 Roadway and Environmental Characteristics

Crashes occurring on roads with a posted speed limit less than or equal to 35 mph were found to decrease the probability of an injury by 0.12. This finding is related to the earlier finding of the multiple vehicle crash severity model, in which speeds above 55 mph increased the probability of an injury, again suggesting that the performance limitations of the heavy truck, such as braking distance or stability, are severely impacted at high speeds.

Another finding from the single vehicle model, similar to the multiple vehicle model, relates to winter roadway conditions. It was found that the probability of an injury in a single vehicle crash on a road with snowy, slushy, or icy surface decreases by 0.11. As was mentioned earlier, this finding is not uncommon in the literature and can be attributed to drivers operating with more caution under adverse roadway or weather conditions.

5.4 Summary

In this chapter, the results of both the single and multiple vehicle heavy truck crash models were discussed in detail. A nested logit model was estimated to examine multiple vehicle crash severity, whereas a binary probit model was estimated to examine single vehicle crash severity. The multiple vehicle crash severity model found considerably more variables to be significant than did the single vehicle model. Keeping this in mind, the multiple vehicle model also achieved a higher goodness of fit. However, both models still managed to encompass a variety of variables, some common between the two models, suggesting that improvements to heavy truck safety can take place on many fronts.

This concludes the discussion of the models developed to examine heavy truck crash severity. The next and final chapter of this thesis summarizes the content of the previous chapters, offers recommendations based on the findings just discussed, discusses the limitations of the present study, and provides suggestions for future work in relation to heavy truck safety.

CHAPTER 6 CONCLUSIONS, LIMITATIONS, AND RECOMMENDATIONS

6.1 Summary and Conclusions

National and state level statistics indicate that the state of Iowa may be experiencing a disproportional share of fatal heavy truck involved crashes with respect to the rest of the country. While several national studies and a few state level studies have investigated heavy truck crashes, no rigorous analysis of heavy truck crashes has been conducted for the state of Iowa. This thesis utilized statewide crash data from 2007 to 2012 to perform an in-depth analysis of heavy truck crashes (23,538 crashes in total) in the state of Iowa, with the goal of gaining insights into the causes, locations and various others factors associated with the severity of heavy truck crashes in Iowa.

The literature review revealed that crash severity can be estimated by employing either ordered or unordered discrete outcome models. In this thesis, heavy truck severity was estimated using unordered discrete outcome models because of the associated flexibility and goodness of fit. Separate models for single and multiple vehicle crashes were estimated. Single vehicle crash severity was estimated using a binary probit model with outcomes of injury (fatal, major, minor, or possible injury) or no injury (PDO), while multiple vehicle crash severity was estimated using a nested logit model with fatal or major injury and minor or possible injury outcomes nested to compensate for their shared unobserved effects. Elasticities and marginal effects were computed to assess the magnitude of the impact of the significant factors on crash severity. The estimation results and implications of the findings from both models are summarized next.

The type of collision involving a heavy truck was found to have a great impact (based on elasticity) on the severity outcome of multiple vehicle crashes. Head-on and broadside crashes were found to increase the probability of an injury while sideswipe crashes were found to increase the probability of no injury. Vehicular rollover had a large effect (based on marginal effect) on the severity outcome of single vehicle crashes. This finding suggests pronounced improvements to truck safety can be made through measures designed to prevent or reduce the severity of a rollover. Countermeasures, with respect to rollovers, recommended by the literature include mandating electronic stability control on all heavy trucks and equipping heavy trucks with side airbags (Perrin et al., 2007).

Time of the day, day of the week, and seasons were all found to have a relationship to multiple vehicle crash severity. Both early morning (5AM to 8AM) and mid-day hours of the day (11AM to 2PM) where found to increase the probability of severe crashes, while late afternoon and early evening hours (3PM to 6PM) were found to increase the probity of no injury crashes. These findings may be of use to law enforcement agencies in developing schedules and establishing enforcement priorities. Further, the severe crashes taking place in the early morning may be attributed driver drowsiness (Barr et al., 2011) and suggest that education measures focused on increasing driver's awareness of their susceptibility to fatigue in the morning could improve heavy truck safety. Crashes at the beginning of the week and over the weekend were also found to increase the probability of a severe crash. These findings too could be used to more efficiently and effectively deploy enforcement efforts. Additionally, the finding of an increase in crash severity toward the beginning of the week supports the finding by Park and Jovanis (2010) that heavy truck drivers tend to be at more risk for a crash after extended off duty times over 46 hours, such as the weekend. Educating drivers to be on alert after extended off duty periods could also improve heavy truck safety. Both models predicted higher probability of injury crashes during the summer and lower probability toward the end of the work week. However, the effect of these variables, in comparison to the other temporal variables discussed, is rather small.

Vehicle characteristics were also found to be associated with crash severity. The elasticity analysis for the multiple vehicle crash severity model showed that indicator variables for frontal impacts generated the highest elasticity with respect to severe crash outcomes, suggesting that improvements in the frontal structures of both heavy trucks, in particular, and non-heavy trucks could impact heavy truck crash safety the most. This effect was also significant but less pronounced in the single vehicle crash severity model. Possible means of improving frontal crash outcomes with heavy trucks, as identified by Perrin et al. (2007), involve the dissipation of crash energy by designing trucks to crush, collapse, and absorb crash energy; or the deflection of crash energy by equipping trucks with impact structures that manage a collision's energy by deflecting the impacting vehicle and reducing the collision energy absorbed by the impacting vehicle.

The type of heavy truck involved in the crash was found to have different effects on the severity outcomes of a single vehicle compared to a multiple vehicle crash. Collisions of combination trucks with other vehicles would increase the severity of multiple vehicle crashes, while single vehicle collisions involving a single unit truck would increase the probability of an injury. This finding suggests that combination trucks potentially pose a greater hazard to the traveling public however exposure should also be factored in before any definitive conclusions are drawn.

Both models found older drivers to be more likely to sustain an injury in crashes involving heavy trucks. This finding is more likely a reflection of the physiological differences between older and younger drivers. Additional information on the associated driving training and experience would help evaluate this finding.

Environmental and roadway factors were also significant in both the multiple and single vehicle crash severity models. Higher posted speed limits increase the probability of an injury in single and multiple vehicle crashes. This is likely related to heavy truck's energy and momentum dynamics, and suggests that improvements in the performance of heavy trucks can greatly influence heavy truck safety. Performance improvements to heavy trucks suggested include forward collision warning systems, collision mitigation braking systems, and lane departure warning systems (Blower and Woodrooffe, 2012). Other heavy truck performance improvements include electronically controlled breaking systems and electronic stability control (Perrin et. al., 2007).

Finally, both models found winter road conditions to decrease the probability of severe crash outcomes. This finding is consistent with past research findings (Lemp et al., 2010; Bham et al., 2012) and is attributable to drivers being more cautious and attentive under such conditions. Moreover, the severity of a multiple vehicle crashes was found to increase during dark, un-lit lighting conditions and decrease under rainfall events. Again these findings are in line with past work (Lemp et al., 2010; Bham et al., 2012; Abdel-Aty, 2003), further validating the results of the models developed.

6.2 Limitations and Recommendations

Care should be taken when interpreting the findings of this study as they are subject to the data used, the assumptions made, and the methodology used. In specific, this thesis examined the most recent six years of crash data for the state of Iowa only. Other areas outside of Iowa wishing to use this information must understand that the findings from this study are specific to Iowa and subject to variability both temporally and spatially. In addition, this thesis did not examine information on the characteristics of the occupants but only collected data on the drivers of the vehicles involved in crashes.

Lastly, the definition of a heavy truck used in this study is another restriction to consider when evaluating the findings from this thesis. The definition of a heavy truck herein was based on vehicle configuration only, not weight or operational restrictions and as such, generalizations drawn from this study apply only to those configurations selected for this analysis.

This research was the first attempt to conduct a heavy truck safety study for the state of Iowa. Though many insights were gained into heavy truck crashes the following recommendations for future research are provided.

1. The review of literature identified driver-related factors to be the cause of a majority of crashes. The data set used did contain some information on driver characteristics, however additional information would be desirable. In particular, it would be desirable to obtain the licensure data of each driver involved in a crash. Attempts to accomplish this were made, but privacy issues, time, and programming constraints stymied these efforts. Obtaining licensure information would allow researchers to better understand how driving experience relates to the crash experience. Licensure information would also facilitate an investigation of licensure restrictions and any correlation between certain restrictions and the occurrence or outcome of a crash.
2. Additional data related to the roadway could be incorporated into a similar study. The data set used contained general roadway information such as posted speed limit, pavement type, and roadway functional class. In future work on heavy truck crashes it may be of benefit to include information on traffic volumes and mix, lane width, number of lanes, median width, shoulder type, or possibly even the embankment adjacent to the road. The Iowa DOT has much of this information in their geocoded Geographic Information Management System (GIMS) database. It is strongly suggested that any further work supplement the crash report data used in this study with the GIMS data.

3. Discussions with motor vehicle enforcement personnel indicated that it would be beneficial to understand how different characteristics of carriers relate to the crash experience. In particular, it would be of interest to investigate the relationship between interstate or intrastate carriers with respect to crash severity or crash frequency. It may also be of interest to see if a relationship between crash frequency or severity and carrier scale of operations, carrier revenues, carrier fleet size, and carrier age exists.
4. Future work may even consider employing different modeling techniques. This thesis used discrete outcome modeling to model crash severity. Future work on crash severity may consider the application of a mixed logit model. Mixed logit models allow parameter values to vary across observations, addressing many weaknesses of the multinomial logit model and facilitating more consistent estimates of parameters and outcome probabilities. However, developing these models is rather time consuming and the interpretation of the results can be challenging.
5. Another future consideration for modeling relates to the level of severity modeled. The present study estimated severity at the crash level as the worst injury sustained by any vehicle occupant involved in a crash. It may be more accurate to estimate crash severity as the most severe injury sustained at the vehicle level instead. This would necessitate modeling heavy trucks and non-heavy trucks, in a collision with a heavy truck, separately, but may provide a clearer understanding of how vehicle- and driver-level factors relate to crash severity.

REFERENCES

Abdel-Aty, M. (2003) Analysis of Driver Injury Severity Levels at Multiple Locations Using Ordered Probit Models. *Journal of Safety Research*, 34(5), 597-603.

Barr, L., Yang, C., Hanowski, R., & Olson, R. (2011) An Assessment of Driver Drowsiness, Distraction, and Performance in a Naturalistic Setting. Retrieved from: http://www.fmcsa.dot.gov/facts-research/research-technology/report/11-010-RP-Distraction-and-Performance.pdf. November 2012.

Ben-Akiva and Lerman, (1985). Discrete Choice Analysis: Theory and Application to Travel Demand. The MIT Press, Cambridge.

Bham, G.H, Bhanu, J.S., & Manepalli, U.R. (2012) Multinomial Logistic Regression Model for Single-Vehicle and Multivehicle Collisions on Urban U.S. Highways in Arkansas. *Journal of Transportation Engineering*, 138(6), 786-797.

Blower, D. & Woodrooffe, J. (2012) Survey of the Status of Truck Safety: Brazil, China, Australia, and the United States. Retrieved from: http://deepblue.lib.umich.edu/bitstream/2027.42/90952/1/102856.pdf. August 2012.

Blower, D. and Kostyniuk, L. (2007) Strategies to Reduce CMV-involved Crashes, Fatalities, and Injuries in Michigan. Retrieved from: http://www.michigan.gov/documents/msp/CMV_Crashes_in_Michigan-FINAL_206130_7.pdf. August 2012.

Blower, D., Green. P., & Matteson, A. (2010) Condition of Trucks and Truck Crash Involvement. *Journal of the Transportation Research Board*, 2194, 21-28.

Cantor, D., Corsi, T., Grimm, C., & Ozpolat, K. (2009) A Driver Focused Truck Crash Prediction Model. *Transportation Research Part E*, 683-693.

Chang, L. & Chein, J. (2012) Analysis of driver injury severity in truck-involved accidents using a non-parametric classification tree model. *Safety Science* 51(1), 17-22.

Chang, L. & Mannering, F. (1999) Analysis of injury severity and vehicle occupancy in truck- and non-truck-involved accidents. *Accident Analysis and Prevention,* 31 (5), 579-592.

Federal Motor Carrier Safety Administration (FMCSA). (2006) Report to Congress on the Large Truck Crash Causation Study. Retrieved from: https://www.fmcsa.dot.gov/facts-research/research-technology/report/ltccs-2006.html. August 2012.

Geedipally, S., Turner, P., & Patil, S. (2011) Analysis of Motorcycle Crashes in Texas with Multinomial Logit Model. *Transportation Research Record: Journal of the Transportation Research Board.*

Giuliano, G., Zhou, J., McFerrin, P., &Miller M. (2009) Commercial Motor Vehicles' Safety-A California Perspective. Institute of Transportation Studies- University of California, Berkley. Retrieved from: http://www.path.berkeley.edu/PATH/Publications/PDF/PWP/2010/PWP-2010-01.pdf. November 2012.

Gkritza, K., Kinzenbaw, C., Hallmark. S., & Hawkins, N. (2010) An empirical analysis of farm vehicle crash injury severities on Iowa's public road system. Accident Analysis and Prevention. 42 (4), 1392-1397.

Harwood, D., Potts, I., Torbic, D., & Glauz, W. (2003) Highway*/Heavy Vehicle Interaction-Synthesis 3*. Washington DC: TRB.

Hughes, R. (2000) Truck Safety in North Carolina: Effectiveness of NCDMV Enforcement Efforts in FY99. Retrieved from: http://www.hsrc.unc.edu/pdf/2000/cvspoo.pdf. August 2012.

Iowa Department of Transportation, *Automatic Traffic Recorder Yearly Report for 2012*. Retrieved from http://www.iowadot.gov/maps//reports(atr)/year2012.pdf#view=fit. April, 2013.

Iowa Department of Transportation, *Iowa 2010 Annual Vehicle Miles of Travel*. Retrieved 11 7, 2012, from http://www.iowadot.gov/maps/msp/vmt/clvmt10.pdf.

Khorashadi, A., Niemeier, D., Shankar, V., & Mannering, D. (2005) Differences in Rural and Urban Driver-Injury Severities in Accidents Involving Large-Trucks: An Exploratory Analysis. *Accident Analysis and Prevention*, 37 (5), 910-921.

Kockelman, K. & Kweon, Y. (2002) Driver Injury Severity: An Application of Ordered Probit Models. *Accident Analysis and Prevention* 34 (4), 313-321.

Lemp, J., Kockleman, K., and Unnikrishnan, A. (2011). Analysis of large truck crash severity using heteroskedastic ordered probit models. *Accident Analysis and Prevention*, 43 (1) 370-380.

Lucke, R. (1999) Uniformity of Roadside Safety Inspections of Commercial Vehicles and Drivers on the National Level. Retrieved from: http://ntl.bts.gov/lib/16000/16600/16692/PB2000104508.pdf. September 2012.

McCartt, A., Cambell, S., Keppler, S., & Lantz, B. (2007) Enforcement and Compliance. *Transportation Research Circular E-C117- The Domain of Truck and Bus Safety Research*, 41-54.

McMurtry, M., Knipling, R., & Harwood, D. (2007) Compatibility of Trucks and Buses with the Roadway Environment. *Transportation Research Circular E-C117- The Domain of Truck and Bus Safety Research*, 138-145.

Moonesinghe, R., Longthorne, A., Shankar, U., Singh, S., Subramanian, R., & Tessmer, J. (2003) An Analysis of Fatal Large Truck Crashes. Retrieved from: http://www-nrd.nhtsa.dot.gov/Pubs/809-569.pdf. October 2012.

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

National Highway Traffic Safety Administration (NHTSA) (2010) *Traffic Safety Facts 2010 Data*. Retrieved from: http://www-nrd.nhtsa.dot.gov/Pubs/811646.pdf. August 2012.

National Highway Traffic Safety Administration (NHTSA) (2010) *Traffic Safety Facts Iowa 2006-2010*. Retrieved from http://www-nrd.nhtsa.dot.gov/departments/nrd-30/ncsa/STSI/19_IA/2010/19_IA_2010.pdf. November 2012.

O'Donnell, C.J. & Connor D.H. (1996) Predicting the Severity of Motor Vehicle Accident Injuries Using Models of Ordered Multiple Choice. *Accident Analysis and Prevention*, 28(6), 739-753.

Park, S. and Jovanis, P. (2010) Hours of Service and Truck Crash Risk. *Journal of the Transportation Research Board*, 2194, 3-10.

Perrin, D., Clarke, R., Knee, H., Kreeb, R., Perel, M., Rau, P., & Svenson, A. (2007) Vehicle Design and Technology. *Transportation Research Circular E-C117- The Domain of Truck and Bus Safety Research*, 123-135.

Potter, J., D. McTiernan, and M. Levasseur. Austroads. (2013) "Improving the Safety of Heavy Vehicles in Urban Areas – A Crash Analysis and Review of Potential Infrastructure and ITS Countermeasures." Retrieved from: https://www.onlinepublications.austroads.com.au/items/AP-R425-13. April 2013.

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

Savolainen, P., Mannering, F., Lord, D., & Quddus, M. (2011) The statistical analysis of highway crash-injury severities: A review and assessment of methodological alternatives. *Accident Analysis and Prevention*, 43 (6), 1666-1676.

Stratham, J., Bronfman, L., & Dong, H. (2010) Improving Commercial Motor Vehicle Safety in Oregon. Retrieved from: http://www.oregon.gov/ODOT/TD/TP_RES/docs/reports/2010/commercialvehiclesafety.pdf. September 2012.

Washington, S., Karlaftis, M., and Mannering, F. (2011). Statisitcal and Econometric Methods for Transportation Data Analysis. CRC Press, Boca Raton, FL.

Xie, Y., Zhoa, K., & Huynh, N. (2012) Analysis of Driver Injury Severity in Rural Single-Vehicle Crashes. *Accident Analysis and Prevention*, 47, 36-44.

APPENDIX A: TESTS OF PROPORTIONS

Test of Proportions: Single and Multiple Vehicle Crash Driver Age

Test of the Difference between Proportions from Two age distributions

2007-2012

341

$$p = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2}$$

$$S_{p_1 - p_2} = \sqrt{\frac{p(1-p)}{n_1} + \frac{p(1-p)}{n_2}}$$

$$z = \frac{p_1 - p_2}{S_{p_1 - p_2}}$$

RELATIVE BRIDGE WIDTH	pop (from licensure data)		crash data (IA driver, vehicle 7-12)	
Total observations	181,314	-	7,803	-
Total Daily VMT		-		-
Total Centerline-Miles		-		-
Weighted Average of ADT		-		-
Crash Rate (per 100M VMT)		-		-
Crash Density (crash per mile)		-		-

Hypothesis being Tested

H_0: $\text{Proportion}_1 = \text{Proportion}_2$

Level of Significance

$\alpha = 0.05$

Driver Age	Frequency	%	Frequency	%	P-Value	Simple Interpretation	Degree of Significance	Number of Observations in Size of Sample 1 n_1	Size of Sample 2 n_2	Proportion 1 p_1	Proportion 2 p_2	Difference between the Sample Proportions $p_1 - p_2$	Weighted Average of p_1 and p_2 p	Estimated Standart Error S_{p1-p2}	z-Stat Value z	Two-tailed probability value p-value
Less than 19	915	0.5%	59	0.8%	0.0024	The two proportions 'are' different.	1	181,314	7,803	0.0050	0.0076	0.0025	0.0052	0.0008	3.0386	0.0024
20-24	7230	4.0%	447	5.7%	0.0000	The two proportions 'are' different.	1	181,314	7,803	0.0399	0.0573	0.0174	0.0406	0.0023	7.6305	0.0000
25-29	11468	6.3%	632	8.1%	0.0000	The two proportions 'are' different.	1	181,314	7,803	0.0632	0.0810	0.0177	0.0640	0.0028	6.2718	0.0000
30-34	14111	7.8%	645	8.3%	0.1190			181,314	7,803	0.0778	0.0827	0.0048	0.0780	0.0031	1.5589	0.1190
35-39	16796	9.3%	715	9.2%	0.7646			181,314	7,803	0.0926	0.0916	0.0010	0.0926	0.0034	0.2994	0.7646
40-44	19489	10.7%	871	11.2%	0.2484			181,314	7,803	0.1075	0.1116	0.0041	0.1077	0.0036	1.1542	0.2484
45-49	23973	13.2%	1035	13.3%	0.9140			181,314	7,803	0.1322	0.1326	0.0004	0.1322	0.0039	0.1080	0.9140
50-54	27264	15.0%	1129	14.5%	0.1689			181,314	7,803	0.1504	0.1447	0.0057	0.1501	0.0041	1.3756	0.1689
55-59	22687	12.5%	948	12.1%	0.3419			181,314	7,803	0.1251	0.1215	0.0036	0.1250	0.0038	0.9504	0.3419
60-64	16681	9.2%	634	8.1%	0.0013	The two proportions 'are' different.	1	181,314	7,803	0.0920	0.0813	0.0107	0.0916	0.0033	3.2239	0.0013
65-69	11363	6.3%	364	4.7%	0.0000	The two proportions 'are' different.	1	181,314	7,803	0.0627	0.0466	0.0160	0.0620	0.0028	5.7459	0.0000
70-74	3861	2.1%	181	2.3%	0.2554			181,314	7,803	0.0213	0.0232	0.0019	0.0214	0.0017	1.1373	0.2554
75-79	3916	2.2%	93	1.2%	0.0000	The two proportions 'are' different.	1	181,314	7,803	0.0216	0.0119	0.0097	0.0212	0.0017	5.8121	0.0000
80-84	1302	0.7%	26	0.3%	0.0001	The two proportions 'are' different.	1	181,314	7,803	0.0072	0.0033	0.0038	0.0070	0.0010	3.9867	0.0001
85-89	245	0.1%	1	0.0%	0.0033	The two proportions 'are' different.	1	181,314	7,803	0.0014	0.0001	0.0012	0.0013	0.0004	2.9351	0.0033
90-94	13	0.0%	0	0.0%	0.4545			181,314	7,803	0.0001	0.0000	0.0001	0.0001	0.0001	0.7480	0.4545
95+	0	0.0%	0	0.0%	#DIV/0!	#DIV/0!	###	181,314	7,803	0.0000	0.0000	0.0000	0.0000	0.0000	#DIV/0!	#DIV/0!
unknown			23	0.3%							0.0029					

(The longer bar indicates the greater proportion value.)

Test of Proportions: Licensure and Crash Data (Gender)

Test of the Difference between Proportions from Two Severity Catagories

2007-2012

341

RELATIVE BRIDGE WIDTH	Multi		Single	
Total Crashes	19,465	-	5,538	-
Total Daily VMT		-		-
Total Centerline-Miles		-		-
Weighted Average of ADT		-		-
Crash Rate (per 100M VMT)		-		-
Crash Density (crash per mile)		-		-

Hypothesis being Tested

H_0: $Proportion_1 = Proportion_2$

Level of Significance

$\alpha = 0.05$

$$p = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2}$$

$$S_{p_1-p_2} = \sqrt{\frac{p(1-p)}{n_1} + \frac{p(1-p)}{n_2}}$$

$$z = \frac{p_1 - p_2}{S_{p_1-p_2}}$$

Driver Age	Frequency	%	Frequency	%	P-Value	Simple Interpretation	Degree of Significance	Number of Observations in Size of Sample 1 n_1	Size of Sample 2 n_2	Proportion 1 p_1	Proportion 2 p_2	Difference between the Sample Proportions $p_1 - p_2$	Weighted Average of p_1 and p_2 p	Estimated Standart Error S_{p1-p2}	z-Stat Value z	Two-tailed probability value p-value
Less than 19	151	0.8%	56	1.0%	! 0.0880	The two proportions 'might be' different	2	19,465	5,538	0.0078	0.0101	0.0024	0.0083	0.0014	1.7061	0.0880
20-24	972	5.0%	342	6.2%	✓ 0.0005	The two proportions 'are' different.	1	19,465	5,538	0.0499	0.0618	0.0118	0.0526	0.0034	3.4779	0.0005
25-29	1,524	7.8%	492	8.9%	✓ 0.0110	The two proportions 'are' different.	1	19,465	5,538	0.0783	0.0888	0.0105	0.0806	0.0041	2.5434	0.0110
30-34	1,585	8.1%	503	9.1%	✓ 0.0257	The two proportions 'are' different.	1	19,465	5,538	0.0814	0.0908	0.0094	0.0835	0.0042	2.2307	0.0257
35-39	1,834	9.4%	565	10.2%	! 0.0820	The two proportions 'might be' different	2	19,465	5,538	0.0942	0.1020	0.0078	0.0959	0.0045	1.7394	0.0820
40-44	2,099	10.8%	672	12.1%	✓ 0.0047	The two proportions 'are' different.	1	19,465	5,538	0.1078	0.1213	0.0135	0.1108	0.0048	2.8256	0.0047
45-49	2,657	13.7%	733	13.2%	✗ 0.4268			19,465	5,538	0.1365	0.1324	0.0041	0.1356	0.0052	0.7946	0.4268
50-54	2,695	13.8%	736	13.3%	✗ 0.2892			19,465	5,538	0.1385	0.1329	0.0056	0.1372	0.0052	1.0598	0.2892
55-59	2,125	10.9%	605	10.9%	✗ 0.9874			19,465	5,538	0.1092	0.1092	0.0001	0.1092	0.0047	0.0158	0.9874
60-64	1,396	7.2%	374	6.8%	✗ 0.2840			19,465	5,538	0.0717	0.0675	0.0042	0.0708	0.0039	1.0714	0.2840
65-69	676	3.5%	199	3.6%	✗ 0.6669			19,465	5,538	0.0347	0.0359	0.0012	0.0350	0.0028	0.4304	0.6669
70-74	322	1.7%	95	1.7%	✗ 0.7538			19,465	5,538	0.0165	0.0172	0.0006	0.0167	0.0020	0.3136	0.7538
75-79	157	0.8%	36	0.7%	✗ 0.2403			19,465	5,538	0.0081	0.0065	0.0016	0.0077	0.0013	1.1743	0.2403
80-84	46	0.2%	17	0.3%	✗ 0.3548			19,465	5,538	0.0024	0.0031	0.0007	0.0025	0.0008	0.9253	0.3548
85-89	7	0.0%	1	0.0%	✗ 0.5110			19,465	5,538	0.0004	0.0002	0.0002	0.0003	0.0003	0.6574	0.5110
90-94	0	0.0%	0	0.0%	#DIV/0!	#DIV/0!	###	19,465	5,538	0.0000	0.0000	0.0000	0.0000	0.0000	#DIV/0!	#DIV/0!
95+	0	0.0%	0	0.0%	#DIV/0!	#DIV/0!	###	19,465	5,538	0.0000	0.0000	0.0000	0.0000	0.0000	#DIV/0!	#DIV/0!
unkown	1,219	6.3%	112	2.0%						0.0626	0.0202	0.0424				

he longer bar indicates the greater proportion valu

Test of Proportions: Multiple Vehicle Crash Severity by Day of the Week

Test of the Difference between Proportions from Two Age Groups

2007-2012

341

RELATIVE BRIDGE WIDTH	Fatal major		Other	
Total Crashes	937	-	17,063	-
Total Daily VMT		-		-
Total Centerline-Miles		-		-
Weighted Average of ADT		-		-
Crash Rate (per 100M VMT)		-		-
Crash Density (crash per mile)		-		-

Hypothesis being Tested

H_0: Proportion$_1$ = Proportion$_2$

Level of Significance

$\alpha = 0.05$

Day of the Week	Frequency	%	Frequency	%	P-Value	Simple Interpretation	Degree of Significance
Sunday	44	4.7%	850	5.0%	0.6951		
Monday	200	21.3%	2,933	17.2%	0.0011	The two proportions 'are' different.	1
Tuesday	167	17.8%	3,113	18.2%	0.7450		
Wednesday	148	15.8%	2,898	17.0%	0.3446		
Thursday	147	15.7%	2,946	17.3%	0.2128		
Friday	136	14.5%	3,040	17.8%	0.0098	The two proportions 'are' different.	
Saturday	95	10.1%	1,283	7.5%	0.0033	The two proportions 'are' different.	

he longer bar indicates the greater proportion valu

$$p = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2}$$

$$S_{p1-p2} = \sqrt{\frac{p(1-p)}{n_1} + \frac{p(1-p)}{n_2}}$$

$$z = \frac{p_1 - p_2}{S_{p1-p2}}$$

Number of Observations in Size of Sample 1	Size of Sample 2	Proportion 1	Proportion 2	Difference between the Sample Proportions	Weighted Average of p_1 and p_2	Estimated Standart Error	z-Stat Value	Two-tailed probability value
n_1	n_2	p_1	p_2	$p_1 - p_2$	p	S_{p1-p2}	z	p-value
937	17,063	0.0470	0.0498	0.0029	0.0497	0.0073	0.3919	0.6951
937	17,063	0.2134	0.1719	0.0416	0.1741	0.0127	3.2664	0.0011
937	17,063	0.1782	0.1824	0.0042	0.1822	0.0130	0.3253	0.7450
937	17,063	0.1580	0.1698	0.0119	0.1692	0.0126	0.9451	0.3446
937	17,063	0.1569	0.1727	0.0158	0.1718	0.0127	1.2459	0.2128
937	17,063	0.1451	0.1782	0.0330	0.1764	0.0128	2.5815	0.0098
937	17,063	0.1014	0.0752	0.0262	0.0766	0.0089	2.9363	0.0033

Test of Proportions: Single and Multiple Vehicle Crashes by Road Surface Conditions

Test of the Difference between Proportions from Two Age Groups

2007-2012

341

$$p = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2}$$

$$S_{p_1-p_2} = \sqrt{\frac{p(1-p)}{n_1} + \frac{p(1-p)}{n_2}}$$

$$z = \frac{p_1 - p_2}{S_{p_1-p_2}}$$

RELATIVE BRIDGE WIDTH	Multi		Single	
Total Crashes	19,465	-	5,462	-
Total Daily VMT		-		-
Total Centerline-Miles		-		-
Weighted Average of ADT		-		-
Crash Rate (per 100M VMT)		-		-
Crash Density (crash per mile)		-		-

Hypothesis being Tested

H_0: $Proportion_1 = Proportion_2$

Level of Significance

$\alpha = 0.05$

Road Surface Conditions	Frequency	%	Frequency	%	P-Value	Simple Interpretation	Degree of Significance	Number of Observations in Size of Sample 1 n_1	Size of Sample 2 n_2	Proportion 1 p_1	Proportion 2 p_2	Difference between the Sample Proportions $p_1 - p_2$	Weighted Average of p_1 and p_2 p	Estimated Standart Error S_{p1-p2}	z-Stat Value z	Two-tailed probability value p-value
Dry	12,495	64.2%	3,009	55.1%	0.0000	The two proportions 'are' different.	1	19,465	5,462	0.6419	0.5509	0.0910	0.6220	0.0074	12.2597	0.0000
Wet	2,271	11.7%	710	13.0%	0.0074	The two proportions 'are' different.	1	19,465	5,462	0.1167	0.1300	0.0133	0.1196	0.0050	2.6805	0.0074
Ice	1,596	8.2%	725	13.3%	0.0000	The two proportions 'are' different.	1	19,465	5,462	0.0820	0.1327	0.0507	0.0931	0.0044	11.4040	0.0000
Snow	2,324	11.9%	409	7.5%	0.0000	The two proportions 'are' different.	1	19,465	5,462	0.1194	0.0749	0.0445	0.1096	0.0048	9.3043	0.0000
Slush	351	1.8%	51	0.9%	0.0000	The two proportions 'are' different.	1	19,465	5,462	0.0180	0.0093	0.0087	0.0161	0.0019	4.5081	0.0000
Sand/Mud/Dirt	255	1.3%	176	3.2%	0.0000	The two proportions 'are' different.	1	19,465	5,462	0.0131	0.0322	0.0191	0.0173	0.0020	9.5805	0.0000
Water	11	0.1%	3	0.1%	0.9651			19,465	5,462	0.0006	0.0005	0.0000	0.0006	0.0004	0.0437	0.9651
Other	57	0.3%	24	0.4%	0.0926	The two proportions 'might be' different.	2	19,465	5,462	0.0029	0.0044	0.0015	0.0032	0.0009	1.6819	0.0926
Unknown	45	0.2%	35	0.6%	0.0000	The two proportions 'are' different.	1	19,465	5,462	0.0023	0.0064	0.0041	0.0032	0.0009	4.7296	0.0000
Not Reported	60	0.3%	320	5.9%	0.0000	The two proportions 'are' different.	1	19,465	5,462	0.0031	0.0586	0.0555	0.0152	0.0019	29.5851	0.0000

he longer bar indicates the greater proportion valu

APPENDIX B: TRANSFERABILITY TEST

LIMDEP output for ALL crashes

```
create; if(x63=1|x64=1)turn=1$
nlogit;lhs=x1;choices=fatmaj,minpos,pdo;model:
u(fatmaj)=dage*x10+speed55*x28+frntimp*x37+rearimp*x39+sprng*x92+bweek*
x96+wknd*x98+AFTRN*x107+rural*x110+UsRt*x115+iart*x116/
u(minpos)=minpos*one+roll*x7+ftyrow*x22+speed552*x28+frntimp2*x37+reari
mp*x39+su*x46+wntrd*x87+summ2*x89+mntuwd*x95+
rural2*x110+UsRt2*x115+iart2*x116+muni*x118+vage*x50+4way*x122/
u(pdo)=pdo*one+roll2*x7+fixed*x9+speed*x21+turn*turn+drkrd*x82+wntrd2*x
87+inrst*x114+muni*x118$

+---------------------------------------------+
| Discrete choice and multinomial logit models|
+---------------------------------------------+
Normal exit from iterations. Exit status=0.
+---------------------------------------------+
| Discrete choice (multinomial logit) model   |
| Maximum Likelihood Estimates                |
| Model estimated: Apr 24, 2013 at 11:33:11AM.|
| Dependent variable               Choice     |
| Weighting variable                 None     |
| Number of observations            21876     |
| Iterations completed                  6     |
| Log likelihood function       -14978.15     |
| Number of parameters                 34     |
| Info. Criterion: AIC =          1.37248     |
|   Finite Sample: AIC =          1.37248     |
| Info. Criterion: BIC =          1.38490     |
| Info. Criterion:HQIC =          1.37652     |
| R2=1-LogL/LogL*  Log-L fncn  R-sqrd  RsqAdj |
| Constants only  -16071.8692  .06805  .06733 |
| Chi-squared[32]          =   2187.43311     |
| Prob [ chi squared > value ] =   .00000     |
| Response data are given as ind. choice.     |
| Number of obs.= 25003, skipped3127 bad obs. |
+---------------------------------------------+
+---------------------------------------------+
| Notes No coefficients=> P(i,j)=1/J(i).      |
|       Constants only => P(i,j) uses ASCs    |
|         only. N(j)/N if fixed choice set.   |
|         N(j)  = total sample frequency for j |
|         N     = total sample frequency.     |
|       These 2 models are simple MNL models. |
|       R-sqrd = 1 - LogL(model)/logL(other)  |
|       RsqAdj=1-[nJ/(nJ-nparm)]*(1-R-sqrd)   |
|         nJ    = sum over i, choice set sizes |
+---------------------------------------------+
```

Variable	Coefficient	Standard Error	b/St.Er.	P[|Z|>z]
DAGE	.00782205	.00229233	3.412	.0006
SPEED55	.84478027	.10191679	8.289	.0000
FRNTIMP	1.19686718	.06442550	18.578	.0000
REARIMP	.10568838	.04942424	2.138	.0325
SPRNG	.10125625	.07435781	1.362	.1733
BWEEK	.30834823	.06641988	4.642	.0000
WKND	.23433829	.09264185	2.530	.0114
AFTRN	.27925171	.07125761	3.919	.0001
RURAL	.58891482	.08652553	6.806	.0000
USRT	.36599309	.07946615	4.606	.0000
IART	.32411656	.09595466	3.378	.0007
MINPOS	2.41666433	.15765070	15.329	.0000
ROLL	.42175488	.10875055	3.878	.0001
FTYROW	.24896227	.06070399	4.101	.0000
SPEED552	.55652092	.05215455	10.671	.0000
FRNTIMP2	.62116061	.03882501	15.999	.0000
SU	.09807956	.03751919	2.614	.0089
WNTRD	.37494964	.09044827	4.145	.0000
SUMM2	.08748127	.04053575	2.158	.0309
MNTUWD	.05242397	.03356239	1.562	.1183
RURAL2	.22167329	.04690014	4.726	.0000
USRT2	.20770310	.04558062	4.557	.0000
IART2	.21454672	.05517812	3.888	.0001
MUNI	.47831250	.15175384	3.152	.0016
VAGE	.00341052	.00221024	1.543	.1228
4WAY	.28248759	.04281989	6.597	.0000
PDO	4.32757282	.15577026	27.782	.0000
ROLL2	-.55138011	.10790485	-5.110	.0000
FIXED	.47716719	.06486579	7.356	.0000
SPEED	.17997639	.08088931	2.225	.0261
TURN	.37461237	.04862788	7.704	.0000
DRKRD	.08542019	.04942761	1.728	.0840
WNTRD2	.69352920	.08549724	8.112	.0000
INRST	.12270759	.04920039	2.494	.0126

LIMDEP output for MULTIPLE vehicle crashes

```
nlogit;lhs=x1;choices=fatmaj,minpos,pdo;model:
u(fatmaj)=dage*x10+speed55*x28+frntimp*x37+rearimp*x39+sprng*x92+bweek*
x96+wknd*x98+AFTRN*x107+rural*x110+UsRt*x115+iart*x116/
u(minpos)=minpos*one+roll*x7+ftyrow*x22+speed552*x28+frntimp2*x37+reari
mp*x39+su*x46+wntrd*x87+summ2*x89+mntuwd*x95+
rural2*x110+UsRt2*x115+iart2*x116+muni*x118+vage*x50+4way*x122/
u(pdo)=pdo*one+roll2*x7+fixed*x9+speed*x21+turn*turn+drkrd*x82+wntrd2*x
87+inrst*x114+muni*x118$

+---------------------------------------------+
| Discrete choice and multinomial logit models|
+---------------------------------------------+
+------------------------------------------------------+
|WARNING:   Bad observations were found in the sample. |
|Found2427 bad observations among   19465 individuals. |
|You can use ;CheckData to get a list of these points. |
+------------------------------------------------------+
Normal exit from iterations. Exit status=0.
+---------------------------------------------+
| Discrete choice (multinomial logit) model   |
| Maximum Likelihood Estimates                |
| Model estimated: Apr 24, 2013 at 11:38:24AM.|
| Dependent variable               Choice     |
| Weighting variable                 None     |
| Number of observations            17038     |
| Iterations completed                  7     |
| Log likelihood function       -11749.33     |
| Number of parameters                 34     |
| Info. Criterion: AIC =          1.38318     |
|   Finite Sample: AIC =          1.38319     |
| Info. Criterion: BIC =          1.39863     |
| Info. Criterion:HQIC =          1.38828     |
| R2=1-LogL/LogL*  Log-L fncn  R-sqrd  RsqAdj |
| Constants only  -12706.2451  .07531  .07439 |
| Chi-squared[32]          =   1913.83593     |
| Prob [ chi squared > value ] =    .00000    |
| Response data are given as ind. choice.     |
| Number of obs.= 19465, skipped2427 bad obs. |
+---------------------------------------------+
+---------------------------------------------+
| Notes No coefficients=> P(i,j)=1/J(i).      |
|       Constants only => P(i,j) uses ASCs    |
|         only. N(j)/N if fixed choice set.   |
|         N(j)  = total sample frequency for j |
|         N     = total sample frequency.     |
|       These 2 models are simple MNL models. |
|       R-sqrd = 1 - LogL(model)/logL(other)  |
|       RsqAdj=1-[nJ/(nJ-nparm)]*(1-R-sqrd)   |
|         nJ    = sum over i, choice set sizes |
+---------------------------------------------+
```

Variable	Coefficient	Standard Error	b/St.Er.	P[|Z|>z]
DAGE	.00703956	.00254287	2.768	.0056
SPEED55	.92803464	.11337259	8.186	.0000
FRNTIMP	1.28550532	.07045688	18.245	.0000
REARIMP	.16152947	.05171668	3.123	.0018
SPRNG	.15405722	.08248711	1.868	.0618
BWEEK	.30467413	.07347887	4.146	.0000
WKND	.26299761	.10369074	2.536	.0112
AFTRN	.25745757	.07792467	3.304	.0010
RURAL	.76041616	.09670787	7.863	.0000
USRT	.36514405	.08794908	4.152	.0000
IART	.31967137	.10601359	3.015	.0026
MINPOS	2.43948322	.17419688	14.004	.0000
ROLL	.06534399	.26896215	.243	.8080
FTYROW	.15506411	.06219583	2.493	.0127
SPEED552	.66233219	.05949877	11.132	.0000
FRNTIMP2	.73920527	.04293792	17.216	.0000
SU	.03099919	.04180645	.741	.4584
WNTRD	.27661171	.09611529	2.878	.0040
SUMM2	.08122422	.04597543	1.767	.0773
MNTUWD	.07578446	.03796686	1.996	.0459
RURAL2	.24906471	.05355651	4.651	.0000
USRT2	.20964373	.05093405	4.116	.0000
IART2	.21288459	.06165044	3.453	.0006
MUNI	.65381025	.17176378	3.806	.0001
VAGE	.00427663	.00247521	1.728	.0840
4WAY	.27504309	.04538101	6.061	.0000
PDO	4.41844865	.17165200	25.741	.0000
ROLL2	-1.40953833	.29983264	-4.701	.0000
FIXED	-.16197837	.18201401	-.890	.3735
SPEED	.13342364	.11133228	1.198	.2308
TURN	.60607473	.08448554	7.174	.0000
DRKRD	-.17926375	.06296281	-2.847	.0044
WNTRD2	.54096659	.09075950	5.960	.0000
INRST	.25574497	.05741294	4.454	.0000

LIMDEP output for SINGLE vehicle crashes

```
nlogit;lhs=x1;choices=fatmaj,minpos,pdo;model:
u(fatmaj)=dage*x10+speed55*x28+frntimp*x37+rearimp*x39+sprng*x92+bweek*
x96+wknd*x98+AFTRN*x107+rural*x110+UsRt*x115+iart*x116/
u(minpos)=minpos*one+roll*x7+ftyrow*x22+speed552*x28+frntimp2*x37+reari
mp*x39+su*x46+wntrd*x87+summ2*x89+mntuwd*x95+
rural2*x110+UsRt2*x115+iart2*x116+muni*x118+vage*x50+4way*x122/
u(pdo)=pdo*one+roll2*x7+fixed*x9+speed*x21+turn*turn+drkrd*x82+wntrd2*x
87+inrst*x114+muni*x118$

+-----------------------------------------------+
| Discrete choice and multinomial logit models|
+-----------------------------------------------+
+-------------------------------------------------------+
|WARNING:   Bad observations were found in the sample. |
|Found 700 bad observations among    5538 individuals. |
|You can use ;CheckData to get a list of these points. |
+-------------------------------------------------------+
Normal exit from iterations. Exit status=0.
+-----------------------------------------------+
| Discrete choice (multinomial logit) model   |
| Maximum Likelihood Estimates                 |
| Model estimated: Apr 24, 2013 at 11:41:04AM.|
| Dependent variable              Choice      |
| Weighting variable                None      |
| Number of observations            4838      |
| Iterations completed                 6      |
| Log likelihood function       -3004.554      |
| Number of parameters                34      |
| Info. Criterion: AIC =          1.25612      |
|   Finite Sample: AIC =          1.25622      |
| Info. Criterion: BIC =          1.30169      |
| Info. Criterion:HQIC =          1.27212      |
| R2=1-LogL/LogL*  Log-L fncn  R-sqrd  RsqAdj |
| Constants only   -3353.4359  .10404  .10088 |
| Chi-squared[32]          =    697.76329      |
| Prob [ chi squared > value ] =    .00000     |
| Response data are given as ind. choice.     |
| Number of obs.=  5538, skipped 700 bad obs. |
+-----------------------------------------------+
+-----------------------------------------------+
| Notes No coefficients=> P(i,j)=1/J(i).      |
|       Constants only => P(i,j) uses ASCs    |
|         only. N(j)/N if fixed choice set.   |
|         N(j) = total sample frequency for j |
|         N    = total sample frequency.      |
|       These 2 models are simple MNL models. |
|       R-sqrd = 1 - LogL(model)/logL(other)  |
|       RsqAdj=1-[nJ/(nJ-nparm)]*(1-R-sqrd)   |
|         nJ   = sum over i, choice set sizes |
+-----------------------------------------------+
```

Variable	Coefficient	Standard Error	b/St.Er.	P[|Z|>z]
DAGE	.00756850	.00544800	1.389	.1648
SPEED55	.50702255	.23607678	2.148	.0317
FRNTIMP	1.23381378	.17423060	7.081	.0000
REARIMP	-.81600675	.22555755	-3.618	.0003
SPRNG	-.07567840	.17751230	-.426	.6699
BWEEK	.37393880	.16136016	2.317	.0205
WKND	.19775664	.21434482	.923	.3562
AFTRN	.23338151	.18247620	1.279	.2009
RURAL	-.17431551	.19297566	-.903	.3664
USRT	.01733669	.20467062	.085	.9325
IART	.12207170	.24382999	.501	.6166
MINPOS	2.33340928	.38687098	6.031	.0000
ROLL	-.08860530	.18113282	-.489	.6247
FTYROW	.77367695	.30305845	2.553	.0107
SPEED552	.27795012	.11192258	2.483	.0130
FRNTIMP2	.23149510	.09706111	2.385	.0171
SU	.32534194	.08632024	3.769	.0002
WNTRD	1.27131937	.32682994	3.890	.0001
SUMM2	.14126425	.08809700	1.604	.1088
MNTUWD	-.00793992	.07426056	-.107	.9149
RURAL2	.12963707	.10187832	1.272	.2032
USRT2	.03455560	.10764410	.321	.7482
IART2	.11502071	.12986550	.886	.3758
MUNI	-.31765133	.33884686	-.937	.3485
VAGE	-.00367255	.00515668	-.712	.4763
4WAY	-.10364007	.15805405	-.656	.5120
PDO	4.41387368	.38753586	11.390	.0000
ROLL2	-1.59730459	.18317743	-8.720	.0000
FIXED	-.06770502	.09908583	-.683	.4944
SPEED	-.06327953	.12754972	-.496	.6198
TURN	.46082738	.11447651	4.026	.0001
DRKRD	.14124487	.08586294	1.645	.1000
WNTRD2	1.79864310	.32008335	5.619	.0000
INRST	-.13245309	.09997901	-1.325	.1852

Calculation:

$$\boldsymbol{likelihood\ ratio = -2[LL(\beta_T) - LL(\beta_a) - LL(\beta_b)]}$$

- **-2[-14,978.15 - $^{-}$11,749.33- $^{-}$3,004.55] = <u>448.54</u>**
- **Degrees of freedom = (34 + 34) – 34 = <u>34</u>**
- **χ^2 critical** *(P=0.95)* = **<u>48.6</u>**
- **448.54 > 48.6, therefore single/multiple vehicle split verified**

APPENDIX C: MULTIPLE VEHICLE MNL MODEL ESTIMATION RESULTS

LIMDEP Output: Multiple Vehicle MNL model

```
nlogit;lhs=x1;choices=fatmaj,minpos,pdo;model:

u(fatmaj)=Dage*x10+HTfrnt*x37+HTrear*x39+comb*combcarg+hdbrd*x69+bweek*
x96+wknd*x98+AFTRN*x107+speed55*x28+dark*x82+
3plus*x133+HTHT*x141+PVdrv60*x165+PVfrn*x185+PVside*pvside+PVage10*x189
+van*x136+car*x134+SUV*x137+PVmultiO*x172+pre*precip/

u(minpos)=minpos*one+HTfrnt2*x37+HTside*htside+hdbrd2*x69+comb2*combcar
g+wintrd2*x87+speed552*x28+dark2*x82+ltsumm*ltsumm+
PVdrv602*x165+fmale*x167+PVfrn2*x185+PVrear*x186+PVage102*x189+3plus*x1
33+PVmultiO*x172/

u(pdo)=pdo*one+sdswi*x68+wintrd*x87+eweek*x97+PVftyrow*x171+octnov*octn
ov+PM*epm$
+-----------------------------------------------+
| Discrete choice and multinomial logit models|
+-----------------------------------------------+
+--------------------------------------------------------+
|WARNING:   Bad observations were found in the sample. |
|Found1857 bad observations among   19465 individuals. |
|You can use ;CheckData to get a list of these points. |
+--------------------------------------------------------+
Normal exit from iterations. Exit status=0.
+-----------------------------------------------+
| Discrete choice (multinomial logit) model   |
| Maximum Likelihood Estimates                |
| Model estimated: Apr 02, 2013 at 11:27:33AM.|
| Dependent variable               Choice     |
| Weighting variable                 None     |
| Number of observations            17608     |
| Iterations completed                  7     |
| Log likelihood function       -11376.62     |
| Number of parameters                 42     |
| Info. Criterion: AIC =          1.29698     |
|   Finite Sample: AIC =          1.29699     |
| Info. Criterion: BIC =          1.31553     |
| Info. Criterion:HQIC =          1.30309     |
| R2=1-LogL/LogL*  Log-L fncn  R-sqrd  RsqAdj |
| Constants only  -13080.8153  .13028  .12924 |
| Chi-squared[40]          =   3408.38195     |
| Prob [ chi squared > value ] =   .00000     |
| Response data are given as ind. choice.     |
| Number of obs.= 19465, skipped1857 bad obs. |
+-----------------------------------------------+
```

```
+-----------------------------------------------+
| Notes No coefficients=> P(i,j)=1/J(i).        |
|       Constants only => P(i,j) uses ASCs      |
|         only. N(j)/N if fixed choice set.     |
|         N(j) = total sample frequency for j   |
|         N    = total sample frequency.        |
|       These 2 models are simple MNL models.   |
|       R-sqrd = 1 - LogL(model)/logL(other)    |
|       RsqAdj=1-[nJ/(nJ-nparm)]*(1-R-sqrd)     |
|         nJ   = sum over i, choice set sizes   |
+-----------------------------------------------+
```

Variable	Coefficient	Standard Error	b/St.Er.	P[|Z|>z]
DAGE	.00555467	.00263701	2.106	.0352
HTFRNT	1.17337787	.08065324	14.548	.0000
HTREAR	.25004579	.11860009	2.108	.0350
COMB	.63482816	.07714891	8.229	.0000
HDBRD	1.48629036	.08449949	17.589	.0000
BWEEK	.30534932	.07678525	3.977	.0001
WKND	.25791782	.10880448	2.370	.0178
AFTRN	.21359715	.08262683	2.585	.0097
SPEED55	1.83286713	.08971381	20.430	.0000
DARK	.38452257	.11310480	3.400	.0007
3PLUS	.66852742	.06299269	10.613	.0000
HTHT	.71238777	.10011349	7.116	.0000
PVDRV60	.16422472	.08401587	1.955	.0506
PVFRN	.97944298	.10476037	9.349	.0000
PVSIDE	.56279187	.09636300	5.840	.0000
PVAGE10	.53739296	.07575304	7.094	.0000
VAN	.40645765	.11495428	3.536	.0004
CAR	.14253931	.08310353	1.715	.0863
SUV	.23407748	.10697864	2.188	.0287
PVMULTIO	.14840411	.04146689	3.579	.0003
PRE	-.45055700	.15761193	-2.859	.0043
MINPOS	3.97114270	.19558714	20.304	.0000
HTFRNT2	.69727514	.06436793	10.833	.0000
HTSIDE	.22324242	.05860323	3.809	.0001
HDBRD2	.50902836	.05398801	9.429	.0000
COMB2	.20446902	.04037504	5.064	.0000
WINTRD2	.44571381	.09858742	4.521	.0000
SPEED552	.85953769	.04275647	20.103	.0000
DARK2	.18262157	.06861968	2.661	.0078
LTSUMM	.10357457	.04741825	2.184	.0289
PVDRV602	-.18019127	.04763315	-3.783	.0002
FMALE	.39527930	.03946686	10.015	.0000
PVFRN2	.37115383	.04766029	7.787	.0000
PVREAR	.12099847	.06181223	1.958	.0503
PVAGE102	.25644335	.04030385	6.363	.0000
PDO	6.17825365	.19159354	32.247	.0000
SDSWI	.75901651	.05345852	14.198	.0000
WINTRD	.86619273	.09516622	9.102	.0000
EWEEK	.07868763	.03950966	1.992	.0464
PVFTYROW	.24311185	.06304075	3.856	.0001
OCTNOV	.15321467	.05352178	2.863	.0042
PM	.16402159	.04640455	3.535	.0004

Multiple Vehicle MNL Model Coefficients

Variable	Fatal or Major Injury		Minor or Possible Injury		PDO	
	Coefficient	*t*-Statistic	Coefficient	*t*-Statistic	Coefficient	*t*-Statistic
Constant						
Minor/Possible Injury Crash	-	-	3.97	20.30	-	-
Property Damage Only (PDO) Crash	-	-	-	-	6.18	32.25
Crash Specific Characteristics						
(HDBRD) Head-on or Broadside Crash	1.49	17.59	0.509	9.429	-	-
(SDSWIPE) Sideswipe (same direction) Crash	-	-	-	-	0.759	14.20
(3PLUS) 3or More Vehicles in a Crash	0.669	10.61	0.669	10.61	-	-
(HTHT) Heavy Truck Crash with Heavy Truck	0.712	7.116	-	-	-	-
(VAN) Crash Involved a Van	0.406	3.536	-	-	-	-
(CAR) Crash Involved a Car	0.143	1.715	-	-	-	-
(SUV) Crash Involved a SUV	0.234	2.188	-	-	-	-
Time and Location Characteristics						
(LTSUMM) Late Summer (July, August, or September)	-	-	0.104	2.184	-	-
(FALL) Fall (October, November)	-	-	-	-	0.153	2.863
(BWEEK) Beginning of the Week (Monday or Tuesday)	0.305	3.977	-	-	-	-
(EWEEK) End of the Week (Thursday or Friday)	-	-	-	-	0.0787	1.992
(WKND) Weekend (Saturday/Sunday)	0.258	2.370	-	-	-	-
(AFTRN) Afternoon (11AM to 2PM)	0.214	2.585	-	-	-	-
(PM) Evening Peak (3PM to 6PM)	-	-	-	-	0.164	3.535
Vehicle Attributes						
(COMB) Cargo Type Combination Truck	0.635	8.229	0.204	5.064	-	-
(HTFRNT) Heavy Truck Front Initial Impact	1.17	14.55	0.697	10.83	-	-
(HTREAR) Heavy Truck Rear Initial Impact	0.250	2.108	-	-	-	-
(HTSIDE)Heavy Truck Side (driveror passenger side) Initial Impact	-	-	0.223	3.809	-	-
(PVFRNT) Passenger Vehicle Front Most Damage	0.979	9.349	0.371	7.787	-	-
(PVSIDE) Passenger Vehicle Side Most Damage (driver or passenger side)	0.563	5.840	-	-	-	-
(PVREAR) Passenger Vehicle Rear Most Damage	-	-	0.121	1.958	-	-
(PVAGE10) Passenger Vehicle 10+ Years Old	0.537	7.094	0.256	6.363	-	-
(PVMULTIO) Passenger Vehicle had Multiple Occupants	0.148	3.579	0.148	3.579	-	-
Driver Characteristics						
(HTAGE) Heavy Truck Driver Age	5.55E-03	2.106	-	-	-	-
(PVDRV60) Passenger Vehicle Driver 60+ Years Old	0.164	1.955	-0.180	-3.783	-	-
(PVFEMALE) Passenger Vehicle Driver is a Female	-	-	0.395	10.02	-	-
(PVFTYROW) Passenger Vehicle Driver FTYROW	-	-	-	-	0.243	3.856
Roadway and Environmental Characteristics						
(SPEED55) Speed Limit 55+ (fatal/major)	1.83	20.43	0.860	20.1	-	-
(WINTRD)Winter Road Surface (Ice, Snow, or Slush)	-	-	0.446	4.521	0.866	9.102
(Precip) Raining or Misting	-0.451	-2.859	-	-	-	-
(Dark) Dark Environment No Road Lighting	0.385	3.400	0.183	2.661	-	-
Log Likelihood at zero			**-13,080.82**			
Log Likelihood at convergence			**-11,376.62**			
Adjusted ρ^2			**0.13**			

Multiple Vehicle MNL Model Elasticities

Variable	Elasticity (%)		
	Fatal/Major Injury	Minor/Possible Injury	PDO
Crash Specific Characteristics			
(HDBRD) Head-on or Broadside Crash	299.3*	-10.0	-10.0
-	-11.7	46.8*	-11.7
(SDSWIPE) Sideswipe (same direction) Crash	-42.7	-42.7	22.4*
(3PLUS) 3or More Vehicles in a Crash	86.6*	-4.4	-4.4
-	-16.5	63.1*	-16.50
(HTHT) Heavy Truck Crash with Heavy Truck	94.2*	-4.7	-4.7
(VAN) Crash Involved a Van	46.2*	-2.6	-2.6
(CAR) Crash Involved a Car	14.4*	-0.8	-0.8
(SUV) Crash Involved a SUV	24.6*	-1.42	-1.42
Time and Location Characteristics			
(LTSUMM) Late Summer (July, August, or September)	-2.4	8.3*	-2.4
(FALL) Fall (October, November)	-10.4	-10.4	4.4*
(BWEEK) Beginning of the Week (Monday or Tuesday)	33.2*	-1.8	-1.8
(EWEEK) End of the Week (Thursday or Friday)	-5.5	-5.5	2.3*
(WKND) Weekend (Saturday/Sunday)	27.4*	-1.6	-1.6
(AFTRN) Afternoon (11AM to 2PM)	22.3*	-1.3	-1.3
(PM) Evening Peak (3PM to 6PM)	-11.1	-11.1	4.7*
Vehicle Attributes			
(COMB) Cargo Type Combination Truck	82.3*	-3.4	-3.4
-	-4.5	17.1*	-4.5
(HTFRNT) Heavy Truck Front Initial Impact	199.0*	-7.2	-7.20
-	-15.7	69.2*	-15.7
(HTREAR) Heavy Truck Rear Initial Impact	26.4*	-1.5	-1.5
(HTSIDE)Heavy Truck Side (driver or passenger side) Initial Impact	-5.0	18.8*	-5.0
(PVFRNT) Passenger Vehicle Front Most Damage	149.8*	-6.10	-6.1
-	-8.4	32.8*	-8.4
(PVSIDE) Passenger Vehicle Side Most Damage	69.3*	-3.6	-3.6
(PVREAR) Passenger Vehicle Rear Most Damage	-3.8	8.5*	-3.8
(PVAGE10) Passenger Vehicle 10+ Years Old	65.2*	-3.4	-3.4
-	-5.7	21.7*	-5.7
(PVMULTIO) Passenger Vehicle had Multiple Occupants	14.9*	-0.9	-0.9
-	-3.4	12.0*	-3.4
Driver Characteristics			
(HTAGE) Heavy Truck Driver Age	0.24*	-0.02	-0.02
(PVDRV60) Passenger Vehicle Driver 60+ Years Old	16.7*	-1.0	-1.0
-	4.1	-13.1*	4.1
(PVFEMALE) Passenger Vehicle Driver is a Female	-8.8	35.4*	-8.8
(PVFTYROW) Passenger Vehicle Driver FTYROW	-16.2	-16.2	6.9*
Environmental Characteristics			
(SPEED55) Speed Limit 55+ (fatal/major)	469.2*	-8.7	-8.7
-	-17.3	95.4*	-17.3
(WINTRD)Winter Road Surface (Ice, Snow, or Slush)	-10.4	39.9*	-10.4
-	-47.2	-47.2	25.5*
(Precip) Raining or Misting	-34.8*	2.3	2.3
(Dark) Dark Environment No Road Lighting	43.4*	-2.4	-2.4
-	-4.3	15.0*	-4.3

*Direct Elasticity

APPENDIX D: SMALL AND HSIAO IIA TEST

Restricted to two outcomes: Fatal/Major or Minor/Possible Outcomes

```
Skip$
probit;LHS=X190;
RHS=one,x9,x36,x38,comb,x68,x95,x97,x106,x27,x164,x184,PVside,x188,x135
precip,x86,x166,x67$

***********************************************************************
 NOTE: Deleted    361 observations with missing data. N is now   5038
***********************************************************************
Normal exit from iterations. Exit status=0.
+---------------------------------------------+
| Binomial Probit Model                        |
| Maximum Likelihood Estimates                 |
| Model estimated: Apr 24, 2013 at 11:04:46AM.|
| Dependent variable                 X190      |
| Weighting variable                 None      |
| Number of observations             5038      |
| Iterations completed                  6      |
| Log likelihood function       -2267.045      |
| Number of parameters                 19      |
| Info. Criterion: AIC =           .90752      |
|   Finite Sample: AIC =           .90755      |
| Info. Criterion: BIC =           .93213      |
| Info. Criterion:HQIC =           .91614      |
| Restricted log likelihood     -2542.230      |
| McFadden Pseudo R-squared      .1082454      |
| Chi squared                    550.3694      |
| Degrees of freedom                   18      |
| Prob[ChiSqd > value] =         .0000000      |
| Hosmer-Lemeshow chi-squared =  28.35804      |
| P-value=  .00041 with deg.fr. =       8      |
+---------------------------------------------+
```

Variable	Coefficient	Standard Error	b/St.Er.	P[|Z|>z]	Mean of X
Index function for probability					
Constant	-2.19297218	.10678425	-20.536	.0000	
X9	.00367870	.00162370	2.266	.0235	45.7328305
X36	.35216931	.04879355	7.218	.0000	.40968638
X38	.19217070	.07262840	2.646	.0081	.14668519
COMB	.28101486	.04500869	6.244	.0000	.55537912
X68	.43086622	.05195053	8.294	.0000	.32949583
X95	.15725226	.04594001	3.423	.0006	.37177451
X97	.23064294	.06689808	3.448	.0006	.11869790
X106	.12169773	.05018540	2.425	.0153	.22866217
X27	.56900314	.04963867	11.463	.0000	.64489877
X164	.14233819	.05095684	2.793	.0052	.22052402
X184	.32087009	.05740004	5.590	.0000	.36780468
PVSIDE	.32289524	.05416209	5.962	.0000	.38725685
X188	.11668572	.04431540	2.633	.0085	.41425169
X135	.14557206	.06739103	2.160	.0308	.10698690
PRECIP	-.22829311	.09174231	-2.488	.0128	.06589917
X86	-.19121289	.05562732	-3.437	.0006	.19392616
X166	-.21764719	.04490340	-4.847	.0000	.43072648
X67	-.27153723	.07076218	-3.837	.0001	.17884081

Fit Measures for Binomial Choice Model Probit model for variable X190		
Proportions P0= .796943		P1= .203057
N = 5038	N0= 4015	N1= 1023
LogL= -2267.045	LogL0= -2542.230	
Estrella = 1-(L/L0)^(-2L0/n) = .10919		
Efron	McFadden	Ben./Lerman
.11751	.10825	.71284
Cramer	Veall/Zim.	Rsqrd_ML
.11338	.19607	.10349
Information Criteria	Akaike I.C.	Schwarz I.C.
	.90752	.93213

Predictions for Binary Choice Model. Predicted value is 1 when probability is greater than .500000, 0 otherwise. Note, column or row total percentages may not sum to 100% because of rounding. Percentages are of full sample.

Actual Value	Predicted Value 0	Predicted Value 1	Total Actual
0	3942 (78.2%)	73 (1.4%)	4015 (79.7%)
1	908 (18.0%)	115 (2.3%)	1023 (20.3%)
Total	4850 (96.3%)	188 (3.7%)	5038 (100.0%)

```
================================================================================
Analysis of Binary Choice Model Predictions Based on Threshold =  .5000
--------------------------------------------------------------------------------
Prediction Success
--------------------------------------------------------------------------------
Sensitivity = actual 1s correctly predicted                              11.241%
Specificity = actual 0s correctly predicted                              98.182%
Positive predictive value = predicted 1s that were actual 1s             61.170%
Negative predictive value = predicted 0s that were actual 0s             81.278%
Correct prediction = actual 1s and 0s correctly predicted                80.528%
--------------------------------------------------------------------------------
Prediction Failure
--------------------------------------------------------------------------------
False pos. for true neg. = actual 0s predicted as 1s                      1.818%
False neg. for true pos. = actual 1s predicted as 0s                     88.759%
False pos. for predicted pos. = predicted 1s actual 0s                   38.830%
False neg. for predicted neg. = predicted 0s actual 1s                   18.722%
False predictions = actual 1s and 0s incorrectly predicted               19.472%
================================================================================
```

Calculation:
***used restricted model from Appendix 5 and full model from Appendix 4A)**

$$\boldsymbol{Small\ and\ Hsiao\ Test} = \frac{\mathbf{1}}{\mathbf{1}-\boldsymbol{N_1}/(\propto\boldsymbol{N})}\{-\mathbf{2}[\boldsymbol{LL}(\boldsymbol{\beta_{UR}}) - \boldsymbol{LL}(\boldsymbol{\beta_R})]\}$$

- $\frac{\mathbf{1}}{\mathbf{1-5{,}038/(1.0\ \mathit{x}\ 17{,}608)}}\{\mathbf{-2[-11,376.62 - -2,267.04]}\}$ **= <u>25,521.32</u>**
- **Degrees of Freedom = <u>19</u>**
- χ^2 **critical** *(P=0.95)* = **<u>30.1</u>**
- **25,521.32 > 30.1, IIA Assumption Violated**

APPENDIX E: LIMDEP OUTPUT: MULTIPLE VEHICLE NESTED LOGIT MODEL

```
nlogit
;lhs=x1
;choices=fatmaj,minpos,pdo
;tree=injury(fatmaj,minpos),noninj(pdo)
;model:
u(pdo)=pdo*one+sdswip*x68+wintrd*x87+eweek*x97+pvftyrow*x171+octnov*octnov
+pm*epm/
u(injury)=hdbrd*x69+pvfrnt*x185+pvage10n*x189+speed55n*x28+pvmultio*x172+3
plus*x133/
u(fatmaj)=dage*x10+HTHT*x141+van*x136+car*x134+suv*x137+bweek*x96+wknd*x98
+aftrn*x107+am*fivtoegh+comb1*combcarg+
precip*precip+HTfrnt*x37+pvside*pvside+dark1*x82+pvdrv60*x165/
u(minpos)=minpos*one+HTfrnt2*x37+comb2*combcarg+fmale*x167+HTSide*htside+w
intrd2*x87+dark2*x82+pvrear*x186+ltsumm*ltsumm
;ivset:(noninj)=[1]
;effects:x10(fatmaj,minpos,pdo)$
+---------------------------------------------+
| FIML Nested Multinomial Logit Model         |
| Maximum Likelihood Estimates                |
| Model estimated: Apr 09, 2013 at 02:05:27PM.|
| Dependent variable                 X1       |
| Weighting variable               None       |
| Number of observations          17608       |
| Iterations completed               65       |
| Log likelihood function     -11542.43       |
| Number of parameters               38       |
| Info. Criterion: AIC =        1.31536       |
|   Finite Sample: AIC =        1.31537       |
| Info. Criterion: BIC =        1.33214       |
| Info. Criterion:HQIC =        1.32088       |
| Restricted log likelihood   -15695.62       |
| McFadden Pseudo R-squared    .2646083       |
| Chi squared                  8306.385       |
| Degrees of freedom                 38       |
| Prob[ChiSqd > value] =       .0000000       |
| R2=1-LogL/LogL*  Log-L fncn  R-sqrd  RsqAdj |
| No coefficients -15695.6248  .26461  .26381 |
| Constants only  -13080.8153  .11761  .11665 |
| At start values -19344.3652  .40332  .40267 |
| Response data are given as ind. choice.     |
+---------------------------------------------+
+---------------------------------------------+
| Notes No coefficients=> P(i,j)=1/J(i).      |
|       Constants only => P(i,j) uses ASCs    |
|         only. N(j)/N if fixed choice set.   |
|         N(j) = total sample frequency for j |
|         N    = total sample frequency.      |
|       These 2 models are simple MNL models. |
|       R-sqrd = 1 - LogL(model)/logL(other)  |
|       RsqAdj=1-[nJ/(nJ-nparm)]*(1-R-sqrd)   |
|         nJ   = sum over i, choice set sizes |
```

```
+---------------------------------------------+
| FIML Nested Multinomial Logit Model         |
| The model has 2 levels.                     |
| Nested Logit form:IV parms = taub|l,r,sl|r  |
| and fr. No normalizations imposed a priori. |
| p(alt=j|b=B,l=L,r=R)=exp[bX_j|BLR]/Sum      |
| p(b=B|l=L,r=R)=exp[aY_B|LR+tauB|LRIVB|LR)]/ |
| Sum. p(l=L|r=R)=exp[cZ_L|R+sL|RIVL|R)]/Sum  |
| p(r=R)=exp[qH_R+fRIVR]/Sum...               |
| Coefs. for branch level begin with HDBRD    |
| Number of obs.= 19465, skipped1857 bad obs. |
+---------------------------------------------+
+---------+--------------+----------------+--------+--------+
|Variable| Coefficient  | Standard Error |b/St.Er.|P[|Z|>z]|
+---------+--------------+----------------+--------+--------+
---------+Attributes in the Utility Functions (beta)
 PDO     |    4.16317279       .37227944    11.183   .0000
 SDSWIP  |     .70891908       .05214510    13.595   .0000
 WINTRD  |     .67747275       .07754976     8.736   .0000
 EWEEK   |     .08711343       .03937872     2.212   .0270
 PVFTYROW|     .24833607       .06252247     3.972   .0001
 OCTNOV  |     .15948472       .05309258     3.004   .0027
 PM      |     .16053589       .04637886     3.461   .0005
 DAGE    |     .00638863       .00265686     2.405   .0162
 HTHT    |     .52641136       .09556560     5.508   .0000
 VAN     |     .59614092       .11480530     5.193   .0000
 CAR     |     .30710611       .08606694     3.568   .0004
 SUV     |     .44339099       .10786737     4.111   .0000
 BWEEK   |     .28171422       .07629682     3.692   .0002
 WKND    |     .28896422       .10999602     2.627   .0086
 AFTRN   |     .23226399       .08441140     2.752   .0059
 AM      |     .21129341       .10860323     1.946   .0517
 COMB1   |     .82349299       .10827929     7.605   .0000
 PRECIP  |    -.44251575       .15741866    -2.811   .0049
 HTFRNT  |    1.36171974       .17807839     7.647   .0000
 PVSIDE  |     .23559038       .07792653     3.023   .0025
 DARK1   |     .48074356       .13328555     3.607   .0003
 PVOLD   |     .26793779       .08317728     3.221   .0013
 MINPOS  |    2.53433928       .16811574    15.075   .0000
 HTFRNT2 |     .87503529       .17272112     5.066   .0000
 COMB2   |     .29022425       .08584523     3.381   .0007
 FMALE   |     .50477210       .06749055     7.479   .0000
 HTSIDE  |     .20904178       .07274476     2.874   .0041
 WINTRD2 |     .32596981       .09663618     3.373   .0007
 DARK2   |     .26510635       .10507521     2.523   .0116
 PVREAR  |     .41785197       .09767614     4.278   .0000
 LTSUMM  |     .11979134       .05992415     1.999   .0456
---------+Attributes of Branch Choice Equations (alpha)
 HDBRD   |     .73979400       .04906960    15.076   .0000
 PVFRNT  |     .44584155       .04514096     9.877   .0000
 PVAGE10N|     .30207680       .03816703     7.915   .0000
 SPEED55N|    1.02981178       .04054738    25.398   .0000
 PVMULTIO|     .14021975       .04141769     3.386   .0007
 3PLUS   |     .65355617       .06301023    10.372   .0000
```

```
---------+IV parameters, tau(b|l,r),sigma(l|r),phi(r)
 INJURY  |      .71112668        .12049826      5.902    .0000
 NONINJ  |     1.00000000    ......(Fixed Parameter).......
```

Calculation

***t*-test $\emptyset_i \neq 1$**

- $t = \frac{\beta - 1}{S.E.(\beta)} = \frac{0.7111-1}{0.1205} = \underline{-2.398}$

- $t_{critical}$ ***(2-tailed, α=0.05)=*** **<u>1.960</u>**

- **\\-2.398\\ >1.96, therefore $\emptyset_i \neq 1$**

APPENDIX F: LIMDEP OUTPUT: SINGLE VEHICLE BINARY PROBIT MODEL

```
Skip$
probit;LHS=X190;
RHS=one,x4,x6,x9,x19,x40,x88,x45,x20,sddmg,x96,x29,x86,anml;marginal
effects;prob=injury2$

 ***********************************************************************
 * NOTE: Deleted    658 observations with missing data. N is now   4804
***********************************************************************
Normal exit from iterations. Exit status=0.
+---------------------------------------------+
| Binomial Probit Model                       |
| Maximum Likelihood Estimates                |
| Model estimated: Apr 18, 2013 at 08:45:32AM.|
| Dependent variable                X190      |
| Weighting variable                None      |
| Number of observations            4804      |
| Iterations completed                 6      |
| Log likelihood function      -2290.026      |
| Number of parameters                14      |
| Info. Criterion: AIC =          .95921      |
|   Finite Sample: AIC =          .95923      |
| Info. Criterion: BIC =          .97809      |
| Info. Criterion:HQIC =          .96584      |
| Restricted log likelihood    -2726.434      |
| McFadden Pseudo R-squared     .1600655      |
| Chi squared                   872.8160      |
| Degrees of freedom                  13      |
| Prob[ChiSqd > value] =        .0000000      |
| Hosmer-Lemeshow chi-squared =   9.60729     |
| P-value=  .29368 with deg.fr. =       8     |
+---------------------------------------------+
+--------+--------------+----------------+--------+--------+----------+
|Variable| Coefficient  | Standard Error |b/St.Er.|P[|Z|>z]| Mean of X|
+--------+--------------+----------------+--------+--------+----------+
---------+Index function for probability
 Constant|   -1.23579372       .09734633   -12.695   .0000
 X4      |     .13014381       .04563460     2.852   .0043    .39696087
 X6      |     .80201324       .04875289    16.451   .0000    .33305579
 X9      |     .00447226       .00158088     2.829   .0047  44.5376769
 X19     |     .41062235       .04841904     8.481   .0000    .34700250
 X40     |     .26402912       .06815102     3.874   .0001    .19046628
 X88     |     .10956590       .05225462     2.097   .0360    .22502082
 X45     |     .23993823       .04858268     4.939   .0000    .24979184
 X20     |     .27371947       .07824297     3.498   .0005    .10387177
 SDDMG   |    -.17488258       .05315835    -3.290   .0010    .54288093
 X96     |    -.09980633       .04652546    -2.145   .0319    .30682764
 X29     |    -.47185235       .06440370    -7.326   .0000    .22647794
 X86     |    -.42299177       .05902627    -7.166   .0000    .23584513
 ANML    |    -.76861844       .13678423    -5.619   .0000    .05495420
```

```
+-----------------------------------------+
| Partial derivatives of E[y] = F[*]   with |
| respect to the vector of characteristics. |
| They are computed at the means of the Xs. |
| Observations used for means are All Obs.  |
+-----------------------------------------+
```

Variable	Coefficient	Standard Error	b/St.Er.	P[|Z|>z]	Elasticity
Index function for probability					
Constant	-.36186383	.02750943	-13.154	.0000	
Marginal effect for dummy variable is P|1 - P|0.					
X4	.03849821	.01362258	2.826	.0047	.07081751
Marginal effect for dummy variable is P|1 - P|0.					
X6	.25465060	.01613110	15.786	.0000	.39301953
X9	.00130956	.00046280	2.830	.0047	.27027515
Marginal effect for dummy variable is P|1 - P|0.					
X19	.12565732	.01537554	8.173	.0000	.20205651
Marginal effect for dummy variable is P|1 - P|0.					
X40	.08205019	.02228799	3.681	.0002	.07241860
Marginal effect for dummy variable is P|1 - P|0.					
X88	.03283031	.01600353	2.051	.0402	.03423341
Marginal effect for dummy variable is P|1 - P|0.					
X45	.07343937	.01544293	4.756	.0000	.08500797
Marginal effect for dummy variable is P|1 - P|0.					
X20	.08662710	.02647410	3.272	.0011	.04169690
Marginal effect for dummy variable is P|1 - P|0.					
SDDMG	-.05148458	.01571899	-3.275	.0011	-.12951922
Marginal effect for dummy variable is P|1 - P|0.					
X96	-.02877585	.01319890	-2.180	.0292	-.04091429
Marginal effect for dummy variable is P|1 - P|0.					
X29	-.12360589	.01470966	-8.403	.0000	-.12972316
Marginal effect for dummy variable is P|1 - P|0.					
X86	-.11256969	.01415221	-7.954	.0000	-.12302711
Marginal effect for dummy variable is P|1 - P|0.					
ANML	-.16321057	.01856571	-8.791	.0000	-.04156250

```
+----------------------------------------+
| Fit Measures for Binomial Choice Model |
| Probit   model for variable X190       |
+----------------------------------------+
| Proportions P0= .745212   P1= .254788  |
| N =    4804 N0=    3580   N1=    1224  |
| LogL=    -2290.026 LogL0=   -2726.434  |
| Estrella = 1-(L/L0)^(-2L0/n) = .17962  |
+----------------------------------------+
|     Efron |  McFadden  |  Ben./Lerman  |
|    .17303 |    .16007  |       .68733  |
|    Cramer | Veall/Zim. |     Rsqrd_ML  |
|    .17608 |    .28921  |       .16614  |
+----------------------------------------+
| Information  Akaike I.C. Schwarz I.C.  |
| Criteria         .95921       .97809   |
+----------------------------------------+
```

```
+-----------------------------------------------------------+
|Predictions for Binary Choice Model.  Predicted value is |
|1 when probability is greater than  .500000, 0 otherwise.|
|Note, column or row total percentages may not sum to     |
|100% because of rounding. Percentages are of full sample.|
+------+---------------------------------+----------------+
|Actual|         Predicted Value         |                |
|Value |       0                1        | Total Actual   |
+------+----------------+----------------+----------------+
|  0   |   3295 ( 68.6%)|    285 (  5.9%)|   3580 ( 74.5%)|
|  1   |    856 ( 17.8%)|    368 (  7.7%)|   1224 ( 25.5%)|
+------+----------------+----------------+----------------+
|Total |   4151 ( 86.4%)|    653 ( 13.6%)|   4804 (100.0%)|
+------+----------------+----------------+----------------+

=======================================================================
Analysis of Binary Choice Model Predictions Based on Threshold =  .5000
-----------------------------------------------------------------------
Prediction Success
-----------------------------------------------------------------------
Sensitivity = actual 1s correctly predicted                     30.065%
Specificity = actual 0s correctly predicted                     92.039%
Positive predictive value = predicted 1s that were actual 1s    56.355%
Negative predictive value = predicted 0s that were actual 0s    79.378%
Correct prediction = actual 1s and 0s correctly predicted       76.249%
-----------------------------------------------------------------------
Prediction Failure
-----------------------------------------------------------------------
False pos. for true neg. = actual 0s predicted as 1s             7.961%
False neg. for true pos. = actual 1s predicted as 0s            69.935%
False pos. for predicted pos. = predicted 1s actual 0s          43.645%
False neg. for predicted neg. = predicted 0s actual 1s          20.622%
False predictions = actual 1s and 0s incorrectly predicted      23.751%
=======================================================================
```

Age Elasticity

```
--> create; elasAGE2=(1-injury2)*x9*(0.00447)$
--> DSTAT;Rhs=ELASAGE2$
Descriptive Statistics
All results based on nonmissing observations.
==========================================================================
=====
Variable     Mean       Std.Dev.     Minimum      Maximum        Cases
Missing
==========================================================================
=====
--------------------------------------------------------------------------
----
All observations in current sample
--------------------------------------------------------------------------
----
ELASAGE2|  .147831      .570515E-01  .253492E-01  .349079         4804
```

Calculation:

$$\boldsymbol{adjusted\ \rho^2 = 1 - \frac{LL(\beta)-k}{LL(0)}} \ = \ \mathbf{1-}\ \frac{\mathbf{-2{,}290.026-14}}{\mathbf{-2{,}726.434}} = \underline{\mathbf{0.155}}$$