Final Report

Reducing Fatalities and Severe Injuries on Florida's High-Speed Multi-Lane Arterial Corridors

Part I:
PRELIMINARY SEVERITY ANALYSIS OF DRIVER CRASH INVOLVEMENTS

BD-548-22

Submitted to
The Florida Department of Transportation
Research Center
605 Suwannee Street, MS 30
Tallahassee FL 32399

Submitted By
Mohamed Abdel-Aty, PhD, PE
Anurag Pande, PhD
Chris Lee, PhD
Abhishek Das
Alexis Nevarez
Ali Darwiche
Premchand Devarasetty

College of Engineering and Computer Science
Center for Advanced Transportation Systems Simulation
University of Central Florida
PO Box 162450
Orlando, FL 32816-2450
Phone: (407) 823-5657
Fax: (407) 823-4676
E-mail: mabdel@mail.ucf.edu

April 2009

1. Report No. **BD-548-22**	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle **Reducing Fatalities and Severe Injuries on Florida's High-Speed Multi-Lane Arterial Corridors;** **<u>Part I:</u> Preliminary Severity Analysis of Driver Crash Involvements**		5. Report Date April 28, 2009 6. Performing Organization Code 1650-7063
7. Author(s) **Mohamed Abdel-Aty, PhD, PE; Anurag Pande, PhD; Chris Lee, PhD; Abhishek Das, Alexis Nevarez, Ali Darwiche, Premchand Devarasetty**		8. Performing Organization Report No.
9. Performing Organization Name and Address **Center for Advanced Transportation and Systems Simulation, University of Central Florida, P.O. Box 162450, Orlando, Fl 32816-2450**		10. Work Unit No. (TRAIS) 11. Contract or Grant No.
12. Sponsoring Agency Name and Address **Florida Department of Transportation 605 Suwannee St. MS 30 Tallahassee, Florida 32399 (850)414-4615**		13. Type of Report and Period Covered **Final Report** 14. Sponsoring Agency Code
15. Supplementary Notes		

16. Abstract

Severe injury involvements on arterial roads account for a quarter of the total severe injuries reported statewide. Crash severity analysis was conducted and consisted of six road entity models and twenty crash type models. The data preparation and sampling was successful in allowing a robust dataset. The overall model was a good candidate for the analysis of driver injury severity on high-speed multilane roads. Driver injury severity resulting from angle and left turn crashes were best modeled by separate unsignalized intersection crash analysis. Injury severity from rear-end and fixed object crashes was best modeled by combined analysis of pure segment and unsignalized intersection crashes. The most important contributing factors found in the overall analysis included driver-related variables such as age, gender, seat belt use, at-fault driver, physical defects and speeding. Crash and vehicle-related contributing factors included driver ejection, collision type (harmful event), contributing cause, type of vehicle and off-roadway crash. Multivehicle crashes and interactions with intersection and off road crashes were also significant. The most significant roadway-related variables included speed limit, *adt* per lane, access class, lane width, roadway curve, sidewalk width, non-high mast lighting density, type of friction course and skid resistance. Two additional models of crashes for urban and rural areas were successfully developed. The land use models' goodness of fit was substantially better than any other combination by road entity or the overall model.

17. Key Word	18. Distribution Statement		
Arterials, Severe crashes, road entity/location, logistic regression			
19. Security Classif. (of this report) **Unclassified**	20. Security Classif. (of this page) **Unclassified**	21. No. of Pages 367	22. Price

EXECUTIVE SUMMARY

Arterial roads constitute the majority of the centerline miles of the Florida State Highway System. Severe injury involvements on these roads account for a quarter of the total severe injuries reported statewide. This research focuses on driver injury severity analysis of statewide high-speed multilane arterials using crash data for the years 2002 to 2004. The first goal is to test different ways of analyzing crash data (by road entity and crash types) and find the best method of driver injury severity analysis. A second goal is to find driver-, vehicle-, road- and environment-related factors that contribute to severe involvements on multilane arterials. Exploratory analysis using one year of crash data (2004) using binary logit regression was used to measure the risk of driver severe injury given that a crash occurs. A preliminary list of significant factors was obtained.

A massive data preparation effort was undertaken and a random sample of multivehicle crashes was selected for final analysis. The final injury severity analysis consisted of six road entity models and twenty crash type models. The data preparation and sampling was successful in allowing a robust dataset. The overall model was a good candidate for the analysis of driver injury severity on high-speed multilane roads. Driver injury severity resulting from angle and left turn crashes were best modeled by separate unsignalized intersection crash analysis. Injury severity from rear-end and fixed object crashes was best modeled by combined analysis of pure segment and unsignalized intersection crashes.

The most important contributing factors found in the overall analysis included driver-related variables such as age, gender, seat belt use, at-fault driver, physical defects and speeding. Crash and vehicle-related contributing factors included driver ejection, collision type (harmful

event), contributing cause, type of vehicle and off-roadway crash. Multivehicle crashes and interactions with intersection and off road crashes were also significant. The most significant roadway-related variables included speed limit, *adt* per lane, access class, lane width, roadway curve, sidewalk width, non-high mast lighting density, type of friction course and skid resistance.

The overall model had a very good fit but some misspecification symptoms appeared due to major differences in road entities and crash types by land use. Two additional models of crashes for urban and rural areas were successfully developed. The land use models' goodness of fit was substantially better than any other combination by road entity or the overall model. Their coefficients were substantially robust and their values agreed with scientific or empirical principles. Additional research is needed to prove these results for crash type models found most reliable by this investigation. A framework for injury severity analysis and safety improvement guidelines based on the results is presented. Additional integration of road characteristics (especially intersection) data is recommended for future research. Also, the use of statistical methods that account for correlation among crashes and locations are suggested for use in future research.

Disclaimer

The opinions, findings, and conclusions expressed in this publication are those of the authors and not necessarily those of the State of Florida Department of Transportation.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS/ABBREVIATIONS

AADT	Annual Average Daily Traffic
AASHTO	American Association of State Highway and Transportation Officials
CAR	Crash Analysis Reporting System
CATSS	Center for Advanced Transportation Systems Simulation
FARS	Fatality Analysis Reporting System
GES	General Estimates System
FDHSMV	Florida Department of Highway Safety and Motor Vehicles
FDOT	Florida Department of Transportation
FHWA	Federal Highway Administration
KABC0	Five level injury scale used in crash reports
NHTSA	National Highway Traffic Safety Administration
RCI	Roadway Characteristics Inventory
SHS	State Highway System
SHSP	Strategic Highway Safety Plan

CHAPTER 1. INTRODUCTION

1.1 Florida State Road Network

The Florida State Highway System (SHS) consists of a complex network of more than 9,700 centerline miles of roads that serve different purposes. This is integrated to a broad inter modal system to provide for the transportation needs of the state residents and visitors. As the main statewide transportation network, it carries a significant amount of vehicular traffic. The arterial roads are an important component of this system, as shown in Table 1-1. More than 95% of the centerline miles of active roads on the SHS serve as arterials. When excluding interstates and expressways, a majority (78.1%) of the SHS roads serve as arterials. When comparing centerline miles totals, the rural portion (43.3%) are higher than the urban (34.5%). However, the lane miles shown in Table 1-2, page 2, show that the built urban capacity (41.1% of the total SHS) is significant higher than the rural capacity (31.9% of the total SHS). These statistics show the importance of the arterials in the SHS and the high degree of urbanized development in the state.

Table 1-1: Distribution of Centerline Miles for Active Roads on the State Highway System (Source: FDOT, 2008)

Functional Classification		Centerline miles			Average Percent
		2002	2003	2004	
Rural	Prin. Arterial - Interstate, toll	1,199	1,199	1,068	9.6%
	Principal Arterial- Other	3,288	3,218	2,821	25.8%
	Minor Arterial	2,140	2,208	2,062	17.7%
	Major Collector	424	424	436	3.6%
	Minor Collector	9	9	0	0.1%
Urban	Prin. Arterial - Interstate, toll	872	872	1,003	7.6%
	Principal Arterial- Other	2,561	2,530	3,009	22.4%
	Minor Arterial	1,422	1,454	1,509	12.1%
	Total Collector	144	138	129	1.1%
Total Active on SHS		12,058	12,051	12,037	100.0%

Table 1-2: Distribution of Lane Miles for Active Roads on the State Highway System (Source: FDOT, 2008)

Functional Classification		Lane miles			Average Percent
		2002	2003	2004	
Rural	Prin. Arterial - Interstate, toll	5,195	5,227	4,672	1232.5%
	Principal Arterial - Other	9,016	8,928	7,715	2095.3%
	Minor Arterial	4,485	4,646	4,300	1096.5%
	Major Collector	877	857	879	213.3%
	Minor Collector	19	19	0	3.1%
Urban	Prin. Arterial - Interstate, toll	4,510	4,547	5,201	1163.6%
	Principal Arterial - Other	11,235	11,264	13,027	2899.1%
	Minor Arterial	4,855	4,989	5,004	1212.0%
	Total Collector	361	352	323	84.6%
Total Active on SHS		40,552	40,828	41,120	10000.0%

The Florida Department of Transportation (FDOT) has engaged in initiatives aimed at improving the balance between access and mobility of the existing and new arterial roads. Two such initiatives are the access management and the corridor management programs, which improve existing road design features to achieve their objectives. The published results suggest that these programs have also been successful at improving safety based on research results (FDOT, 2007 and Williams, 2004). The FDOT is currently investigating how to apply the corridor concept to safety evaluations of high-speed multilane arterial corridors. Nationwide, different *safe corridor* initiatives have been started in states such as New Jersey, Washington, Virginia and Ohio. The concept of corridor safety is an attempt to analyze crash experience along a corridor considering the entire roadway as one entity rather than dividing the crashes into the traditional road entities (segments and intersections) and treat them as *isolated locations*. The traffic behavior and close proximity of some of these road entities on high-speed multilane arterials may affect the crash experience at neighboring locations.

The importance of preventing severe crashes has been acknowledged by government officials. The first of the Safety Recommendations in AASHTO's Surface Transportation Policy

Recommendations for the National Surface Transportation Policy and Revenue Study Commission states (AASHTO, 2007): "Establish a presidential commission to assist in the development of a national strategic highway safety plan designed to drive down fatal and disabling injuries on the nation's highways." Following a similar policy, Florida's Strategic Highway Safety Plan (SHSP) goal states (FDOT, 2006): "To improve the safety of Florida's surface transportation system by achieving a five percent annual reduction in the rate of fatalities and serious injuries beginning in 2007." This underscores the importance of analyzing severe crashes as a group, as opposed to fatal crashes only.

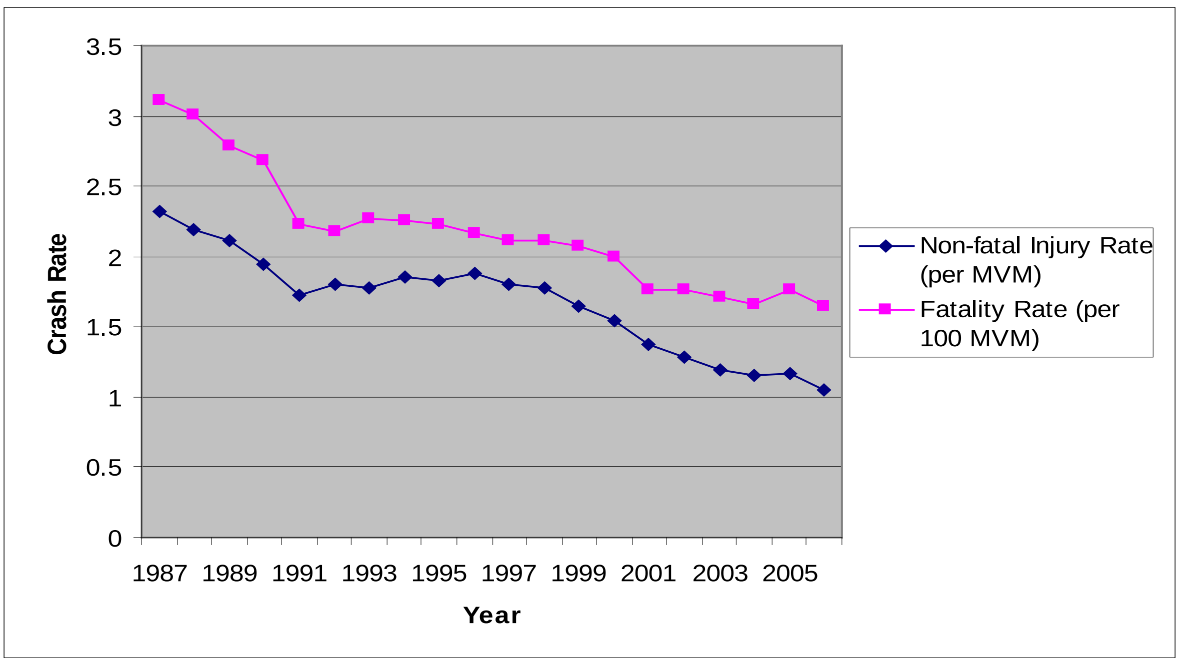


Figure 1-1: Injury Severity Rates per Vehicle Miles Traveled from 1987-2006 (Source: FDHSMV, 2006)

In Florida there has been progress in achieving crash rates reductions in both injury and fatal crashes, as shown in Figure 1-1. However, since 2001 the rates of non-fatal injuries (levels 2-4 in the Florida injury scale) kept a nearly constant slope while the fatality rates have a significantly less pronounced downward trend. Perhaps some of these earlier positive effects are being negated by increased development around arterial corridors in rural areas and riskier driver behavior in the case of the fatalities, among other factors. The difference between the fatal and disabling crash injury outcomes for a given crash will largely depend on driver characteristics when controlling for other factors. The importance of considering the contributing factors of severe crashes in order to prevent fatal injuries is paramount.

Table 1-3: Difference from Previous Year in Involvement Rates per Vehicle Miles Traveled from 2002-2006 (Source: FDHSMV, 2006)

Year	Non- Fatal Injury Rate (per 100 million *vmt*)	Percent Difference	Fatality Rate (per 100 MVM)	Difference with previous year
2002	128.504	-6.32%	1.759011	-0.15%
2003	119.3905	-7.09%	1.712436	-2.65%
2004	115.4889	-3.27%	1.655636	-3.32%
2005	116.3981	0.79%	1.757939	6.18%
2006	105.4622	-9.40%	1.651266	-6.07%
Average		**-5.06%**		**-1.20%**

To consider the trends in recent years, Table 1-3 shows that the average percent difference for the last five years was -5.06 for non-fatal injuries vs. -1.20 for fatalities. Improvements in vehicles and roads may have had some impact on the reduction of injuries; however, there is not a clear trend of reductions in the last five years. Additional measures are required in order to achieve the goal of the 2006 SHSP in a regular basis, especially on the fatal injuries. The non-fatal injury rate groups serious and non-serious injuries, which does not show the individual trend of serious injuries. Analysis can be improved by grouping serious (incapacitating) and fatal injury because similar crash circumstances can lead to either injury outcome. This study focuses on a road user group to which the severe injuries can be directly compared to the vehicular traffic, the drivers.

1.2 Driver Severe Injuries in Florida

Driver involvements have been regarded as an important indicator of the vehicular crash outcomes. The analysis shown in Table 1-4 compares the total severe injuries reported and the driver injuries extracted from the FDHSMV crash database. The drivers (including bicyclists and motorcyclists) represent more than 65% of the total severe injuries, thus their importance in considering contributing factors for severe injury outcomes. The importance of the arterial corridors in the safety picture can be better understood by comparing all of the driver involvements vs. those occurring at high-speed multilane arterials. It can be seen that the percent of severe injury involvements are slightly higher for these arterial corridors (see Table 1-5, page 6). If we consider that minor crash data are underreported in this crash database (Abdel-Aty and Keller, 2005) the true proportion may not be easily obtainable, but it should follow a similar trend for most road types.

Table 1-4: Severe (Incapacitating and Fatal) Injuries from 2004-2006 (*Source: FDHSMV, 2006)

Year	Total severe injuries*	Driver severe injuries	Percent Driver
2002	33,664	22,194	65.93%
2003	32,771	21,800	66.52%
2004	32,792	21,784	66.43%

Table 1-5: Severe (Incapacitating and Fatal) Injuries Sustained by Drivers from 2004-2006

All involvements					Involvements on high-speed multilane roads				
Year	Non-	Severe	Total	Percent	Year	Non-	Severe	Total	Percent

	severe			Severe
2002	383033	22194	405227	5.48%
2003	375171	21800	396971	5.49%
2004	395861	21784	417645	5.22%
Total	1154065	65778	1219843	5.39%

	severe			Severe
2002	93626	5491	99117	5.54%
2003	88924	5218	94142	5.54%
2004	92593	5138	97731	5.26%
Total	275143	15847	290990	5.45%

Another measure of the safety performance of state roads classified as arterial (non limited access) in Florida can be obtained by comparing the driver severe injuries occurring on these corridors in relation to the portion of the total public road composition in the state. Severe driver injuries at high-speed multilane corridors (15,847 out of 65,778) represent 19.41% of the total severe driver injuries in all of Florida's public roads (see Table 1-5). This a significant number in itself, justifying special analysis of the characteristics of severe injury crashes in order to reduce the rates of severe injury. This need is more compelling when we consider the fact that the types of road targeted in this study represent less than 8% of the total centerline miles of all the public roads in Florida (FDOT, 2008). The impact of improving safety conditions in these types of roads has a great potential in improving the safety performance of Florida's road network.

Table 1-6: Severe (Incapacitating and Fatal) Injury Rates for Drivers from 2002-2004

All driver involvements in Florida				
Year	**Annual *vmt* (millions)**	**Severe injuries**	**Severe injury rate (per 100 million *vmt*)**	**Difference with previous year**
2002	178,681	22,194	12.42	--
2003	185,511	21,800	11.75	-5.4%
2004	196,444	21,784	11.09	-5.6%
Drivers on state arterials (non limited access)				
Year	**Annual *vmt* (millions)**	**Severe injuries**	**Severe injury rate (per 100 million *vmt*)**	**Difference with previous year**
2002	58,279	5,491	9.42	--
2003	59,648	5,218	8.75	-7.2%
2004	60,328	5,138	8.52	-2.6%

In terms of severe injury rates, drivers have experienced lower rates in the recent years, but these rates have not been reduced consistently on the state arterial corridors (see Table 1-6). Even when the injury rate is lower than the overall rate, this analysis is focused on the potential for improvement, which is certainly higher for severe crashes than for fatal crashes only (refer to Table 1-3, page 4). These statistics serve to illustrate how the severe injury group analysis helps to better understand the safety performance of the highway network. Additional analysis by road entities and crash types will improve the understanding of injury severity outcomes at high-speed multilane roads.

1.3 Research Objectives

This research focuses on a driver injury severity analysis of statewide high-speed multilane arterials using crash data for the years 2002 to 2004. These were defined as non-limited access roads with four or more lanes and speed limits greater than or equal to 40 mph under state jurisdiction. An additional effort was made to merge data from different databases to include driver, crash, road and environmental factors not usually found in the crash report. Adequate sample size was found to concentrate the statistical modeling on the most severe (incapacitating injury and fatal injury) crashes. Previous research has shown that the accuracy of crash report information increases with the injury level (Hauer and Hakkert, 1988; Elvik and Mysen, 1999). This analysis, due to its systematic nature, will rely almost exclusively on the crash database information and therefore the most accurate data should be pursued.

The first goal of this work is to test different ways of analyzing crash data and find the best combination of road entities applicable to driver injury severity analysis of high-speed multilane arterials. The second goal is to find a group of driver, vehicle, road and environmental factors that contribute to the occurrence of severe crash involvements on multilane arterials in Florida. These objectives are to be achieved considering crash risk effects on the arterial road network as a whole. In order to achieve this, a series of steps were designed as detailed in Section 1.4.

1.4 Research Steps

A series of steps were taken to find the factors affecting severe crash involvements related to Florida's high-speed multilane arterials. The results of a series of severity analyses are used as a guide to better understand the nature of the complex relationships between driver behavior, road features and environmental characteristics and severe crashes on the particular group of roads under study. By selecting the driver injury severity analysis as the fundamental response, we have some major benefits. First, severity analysis has the distinct advantage that by looking at the driver involvements we capture very important crash injury contributing factors not available in aggregate frequency analysis. Second, driver involvement injury severity provides an unbiased measure for injury exposure for crashes involving at least one vehicle, i.e. each vehicle has one driver. Third, driver-related factors that affect severe injury crashes have been recognized to have a great influence in the occurrence of crashes.

The research included two datasets of analysis (see Figure 1-2, page 10): an exploratory analysis using only one year of crash data and the final analysis using three years of crash data. The regression analysis for both datasets followed a data subset methodology explained in Section 3.6. The results of the exploratory analysis were used in an improved data preparation and variable setup for a final analysis using three years of crash data. In both, the preliminary analyses consisted of bivariate statistical analysis between categorical, continuous variables and the driver injury severity variable. The regression analysis included several driver injury severity models in one or two main categories: road entity and crash type. The analysis of the results included a comparison of the models as well as an examination of the effects of the contributing factors for each model. With this process, the main research goals were achieved.

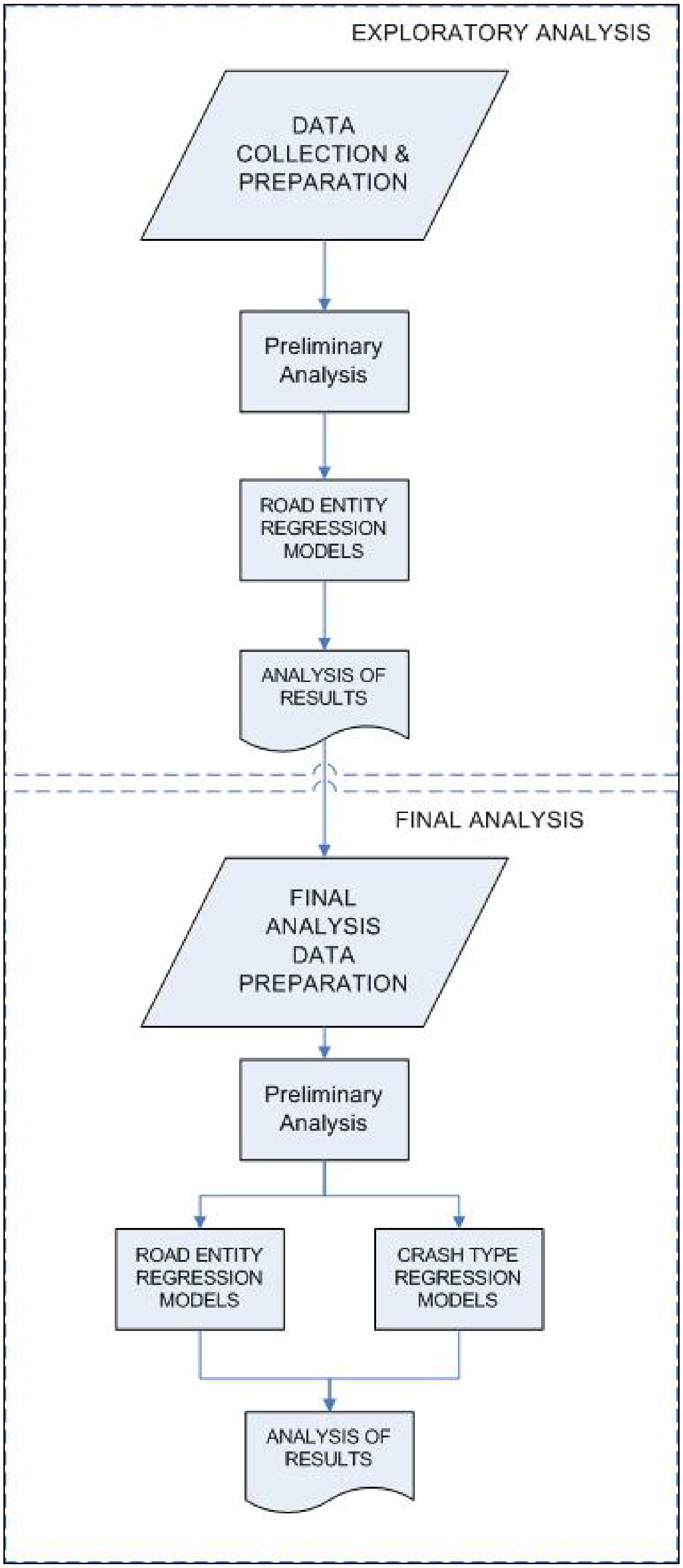


Figure 1-2: Overview of Research Steps

The scope of work is not all inclusive in regards to the marginal contributions of the different factors to injury severity of different drivers involved in a crash. The main objectives of this research are accomplished by considering all the complete records available within the two main clusters of drivers. This issue will be addressed in the next chapters. The results of this research will serve as a foundation for a more advanced systematic analysis of high-speed arterial corridors that is being undertaken for the first time using data from the state of Florida. This research contributes to the body of knowledge by modeling driver injury severity for a sample across a large jurisdiction. Also by comparing different crash types and road entities, the comparison of the reliability of the models are useful in determining future modeling strategies.

CHAPTER 2. LITERATURE REVIEW

2.1 Driver Injury Severity Analysis

2.1.1 Introduction

The modeling of traffic crash injury severity using statistical techniques has been employed as a powerful means to assess road safety conditions. Several situations have warranted the use of this analysis in the past and continue to do so. Some of the uses for injury severity analysis include:

- For a before and after crash severity analysis in order to measure the effectiveness of certain countermeasure in reducing crash severity.
- For conducting a multivariate statistical analysis of crash injury severity based on historical data in order to find contributing factors for certain injury outcomes.
- As part of a systematic study to assess the safety conditions in a large road network.
- When researchers aim to investigate the relationships between driver (and/or vehicle) characteristics and the crash outcomes.

Driver injury severity analysis is a commonly used method to find contributing factors to severe crashes and covers a broad range of crash situations. In Florida, motorcyclists and bicyclists are also considered drivers in the crash reports, thus including groups that have a higher risk of severe crashes. There are challenges when examining different groups in a statistical model. To consider drivers as the unit of analysis has several advantages, such as:

- Severity analysis has the distinct advantage that by looking at the driver involvements we capture very important crash injury contributing factors not available in aggregate frequency analysis.
- Driver involvement injury severity provides an unbiased measure for injury exposure for crashes involving at least one vehicle, i.e. vehicles have each one driver.
- Driver-related factors that affect severe injury crashes have been recognized to have a great influence in the occurrence of crashes.

2.1.2 Definition of Injury Severity Levels

The most commonly used means by police officers to classify the injury severity of persons resulting from a traffic crash is a five level scale, which may vary in definition by jurisdiction. Other injury scales are used in the medical field, but are not currently available in the crash reports used in this analysis. Most state jurisdictions in the United States use the KABC0 five level injury scale: fatal (K), incapacitating injury (A), non-incapacitating (B), possible injury (C), no injury (0). In Florida, an equivalent scale is used with numbers instead of letters. The crash reports' information collected from the FDHSMV traffic crash database (FDHSMV, 2004) follows a five-level injury severity scale, and is defined/numbered as follows:

1) No injury – Indicates there is [no] reason to believe any person received bodily harm from the crash. (Also known as property damage only or PDO)
2) Possible injury – No visible signs of injury but complaint of pain or momentary unconsciousness.
3) Non-incapacitating Evident Injury – Visible injuries from the crash such as bruises, abrasions, limping, etc.

4) Incapacitating injury – Any visible signs of injury from the crash and person(s) had to be carried from the scene.
5) Fatal Injury – An injury sustained in a crash [that] results in a death within 30 days of the crash.

Injury severity analysis has been a mainstay of recent traffic safety research literature due to its intrinsic value to predict factors that influence the main crash outcome (personal injury). The main purpose of this analysis is to describe the relationships that affect crashes with different levels of injury. There are different strategies that can be used to fulfill this goal. One of the strategies is to compute the relative (or conditional) probability of a severe injury given that a crash occurs. The advantage of this type of research is the ease of computation and interpretation of results compared to other methods. In this research the response variable is whether the driver suffered a severe crash (defined as injury levels 4 and 5) given that a crash occurs. A review of the previous studies involving injury severity analysis follows.

2.2 Past Studies Related to Injury Severity Analysis

2.2.1 Relationships between Driver, Vehicle, Traffic and Road Factors and Severity

An exhaustive literature review showed how the injury severity (crash outcome) analysis has evolved, from a tool used to establish relationships between driver factors and the severity of a given crash to a more comprehensive traffic safety analysis tool used for a variety of purposes. Some of the most important objectives that are mentioned next include predicting certain driver's group risk of severe crashes, find certain roadway and environmental characteristics that can be linked to higher crash severity, compare the risk of severe crashes under certain conditions in different geographical regions, and to assess the tradeoffs of reducing certain crash types, which tend to be severe, on the overall crash risk. Many statistical analysis techniques have been employed, usually a type of parametric regression. Also, research in other fields has allowed development of techniques to compare the reliability of different statistical analysis techniques.

Kim et al. (1995) found a relationship between driver characteristics, behavior, vehicles types, crash types, and driver injury severity using crash data in the state of Hawaii during the year 1990. The injury severity log-linear models were developed categorizing by three levels of crash types: head-on, rollover and others. The injury severity response variable was collapsed into four levels: none, possible or non-incapacitating, incapacitating, and fatal. After computing conditional odds derived from a log-linear model, they found that driver behavior (alcohol or drug use and lack of seat belt use) greatly increase the odds of more severe injuries. Driver errors were found to have a small effect, while driver characteristics (age and gender) were mostly insignificant. The main contribution of this and other early work was to demonstrate that there is a complex relationship between driver characteristics, behavior, crash type and injury severity (see Figure 2-1, page 16).

A follow-up study by Richardson et al. (1996) used crash data from Hawaii for the years 1991-1992; they examined the relationships and possible interactions between driver's age, gender, crash type and vehicle type. First, by using categorical data analysis, which guided the development of log-linear models main effect relationships, two- and three-way interactions were found to be significant. The authors found that young drivers have much greater frequency of rollovers and at-fault in rear-end and head-on crashes. Meanwhile, older drivers have much higher frequency of being rear-ended or side-swiped. These findings showed a tendency of different driver groups to be involved in different crash types, suggesting different abilities and behaviors. In addition, it was noted that the reported seat belt use of 97% in crash reports was larger than the observed rate of 85%. Due to the mandatory seat belt law in Hawaii, drivers are motivated to report to the police officer that they were using seat belts at the time of the crash. However, for critically injured or fatal crashes the seat belt use rate drops (below 50% for killed). This apparent over reporting of seat belt use in the lower severity crashes may falsely increase the effectiveness of seat belt use in the model. The authors conducted a sensitivity analysis and found a small effect compared to the total differences in rates across the injury categories.

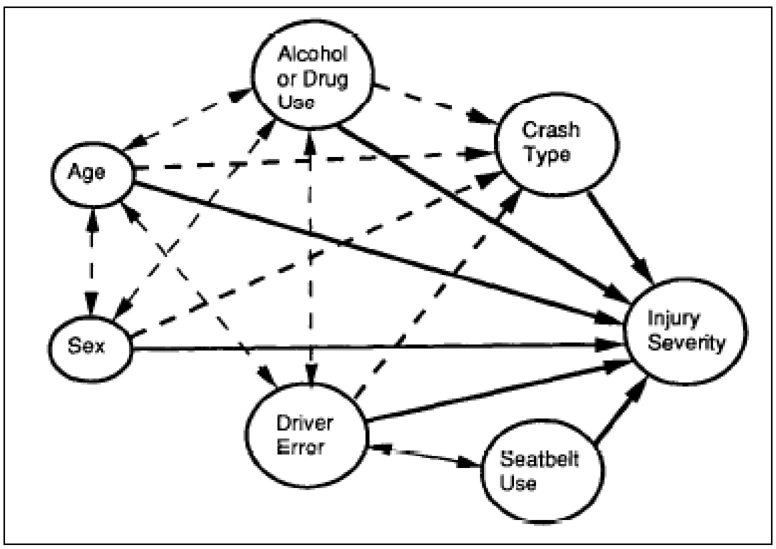


Figure 2-1: Complete Structural Model for Driver Injury Severity (Source: Kim et al., 1995)

By using categorical analysis techniques, including log-linear models, Abdel-Aty et al. (1998) found a broader set of relationships between driver behavior, crash types, traffic, road characteristics and driver injury severity among various driver age groups. Using Florida crash data for the years 1994-1995, four hierarchical models with up to two-way interactions were developed. After computing odd multipliers, the results indicated significant relationships between the driver age and *adt*, injury severity, manner of collision, speed, alcohol involvement and roadway character. The model that included the age, injury severity and *adt* factors suggested that injury severity is related to age and that those old and very old drivers are more likely to be killed in crashes.

When modeling injury severity two main units of analysis are utilized. An aggregate unit of analysis is the crash severity defined as the most serious severity of any of the persons involved in this event. This type of analysis is most useful for larger jurisdictions and when the researcher is more interested in aggregate measures, such as traffic, road types, and environmental factors that contribute to serious crashes or it could also be applied when analyzing single vehicle crashes only. On the other hand, when the researcher whishes to find other important contributing factors related to each driver and vehicle involved in the crash, the person(s) involved becomes the unit of analysis. This disaggregated method is also called analysis of involvements and while more challenging; it provides additional information not available in crash severity analysis. An overview of both methods is presented in the next two sections.

2.2.2 Crash Injury Severity Regression Analysis

Using data from a rural freeway in Washington State, Shankar et al. (1996) found significant effects of environmental conditions, highway design, accident type, driver characteristics and vehicle attributes on crash severity. Crash data from a 61 km study section of a rural interstate during a five-year period (1988-1993) included 1505 single vehicle crashes used in the analysis. A nested logit formulation was used to determine crash severity risk given that a crash occurred. Some of the most important effects found significant were crash types, speeding, restraint use, occupant ejection, driver gender, roadway curves, weather (snow), vehicle-mass (vehicle type) difference and off-road crash location. Interactions were also found between age and sobriety, curves and sobriety, nighttime and icy pavement, fixed object and icy pavement and fixed object and curves. It was noted that the uncertainty about restraint use for less severe crashes has a minimal significance in the models, while more severe crashes reports have more reliable restraint use information. Also, underreporting of the less severe crashes may have a small effect on the model, but the coefficients should continue to be unbiased.

One of the few studies that focused exclusively on environmental effects investigated the major contributing factors to crash severity for one road (State Route 3) in Washington (Lee and Mannering, 2002). This road had rural and urban sections, although model information was shown for the rural sections only. A nested logit model of the severity of 489 run-off road crashes was developed. The five severity levels were combined into three: no evident injury, evident injury and disabling injury/fatality. The results of this model and its marginal effects indicated that the most important factors affecting crash severity included wet road, high-speed road, guardrail location, speeding indicator, asphalt shoulder and weekend indicators. Although the data for this study are limited and thus no general conclusions could be made, some of these environmental factors have been found significant in broader studies. A contribution of this study was the application of the severity models and the marginal effects of the contributing factors as an effective way to compare and analyze the effect of possible countermeasures to run-off road crashes.

A study on single vehicle crashes in an urban area (Hong Kong) investigated the likelihood of fatal or serious injury crashes vs. crashes in which no person involved sustained fatal or serious injury (Yau, 2004). Logistic regression was the statistical tool chosen due to the binary nature of the response variable. Statistical independence test (using contingency tables) and the Cramer's V measure of association were used to guide the variable selection process. The analysis and modeling was developed by type of vehicle (private, commercial, motorcycle) and the risk factors were compared. For private vehicles, district jurisdiction, gender of driver, age of vehicle, time of day and street light conditions were found significant factors. For commercial (goods) vehicles seat-belt usage and weekday occurrence were the only two significant factors, while for motorcycles age of vehicle, weekday and time of day were found to be significant. The author also explained that modeling by vehicle type has the advantage of reducing the heterogeneity in the data, while a disadvantage was the reduction in sample size (less than N=1,000 for each model). To balance the loss in statistical power, a 10% level of significance (entry and exit) was used in the stepwise method.

An unusual (albeit valid) application of crash injury severity analysis was developed by Obeng (2007). Data for a 45 month period from January 2000 to September 2003 for 303 signalized intersections in the city of Greensboro, North Carolina were collected and merged for the analysis. This included intersection characteristics and crash data that were analyzed month by month for each intersection. Months in which the intersection crash totals equal zero were excluded. After the data preparation process, an unbalanced panel data (in terms of months and intersections) resulted in 4,767 crash observations. The dependent variable (possible injury vs. any other injury) was chosen out of necessity of the local jurisdiction: about 85% of the crashes with injuries were classified as possible injuries. This distinction is most useful for the purposes of this research. A binomial logit model was developed and resulted in a good discrimination (85.17% percent correct). Significant factors included driver gender, occupant, number of vehicles involved, driver condition, airbag use, vehicle type, residential land use, sidewalk, amber time, speed limit, a month indicator and interactions of the log *adt*/lane with sidewalk and solid median with pedestrian signal. This and other models involving signalized intersections found a variety of local intersection characteristics significant to the injury risk.

2.2.3 Involved Person Injury Severity Regression Analysis

On of the early comprehensive studies of road crash victims used data from the New South Wales, Australia Road and Traffic Authority (O'Donnell and Connor, 1996). A total of 18,069 motor vehicle occupant involvements occurred during 1991 (representing a census) were used in the analysis. A total of 11 road user attributes were introduced for model fitting. Two modeling types were developed: ordered probit (normal) and ordered logit (logistic) due to the unknown nature of the error distribution. Four injury levels were included in the response: uninjured (also labeled as non-treated injury), treated injury (by a doctor, nurse, or paramedic), admitted injury (hospital), and death (within 30 days and attributed to the crash). The continuous variables age of occupant, speed of vehicle, age of vehicle and time of crash were scaled dividing by 100. The variable selection process utilized the Schwarz Bayesian Information Criterion (SBIC) best subsets method. Significant factors included age of the occupant, vehicle speed, seating position, blood alcohol level, age of vehicle, vehicle type, vehicle make, seat belt usage, and type of collision

A study of age and gender as predictors of driver and front seat passenger injury severity (Mercier et al., 1997) used crash data from Iowa between 1986 and 1993. The data selected for analysis included head-on crashes on high-speed (55 to 65 mph) roads with fatal, major, minor injuries of the driver and front right seat occupant (2,171 injury observations in total). The main objective of the research was to find whether age and gender influenced the severity of the head-on collisions. Results of logistic regression analysis indicated that four separate factors (age, gender, safety equipment, and position in the vehicle) with a total of 14 individual and interactive variables were significant. When the population was divided by gender it was found that age remained significant while seat belts were more beneficial for men and air bags seemed to be more beneficial for women. However, the air bags deployed sample (35 observations) was too small to make a strong conclusion.

To describe the factors that influence the injury severity of passenger vehicle occupants in car-truck rear-end collisions on divided roads, crash data from North Carolina for the years 1993-1995 were used (Duncan et al., 1998). To have a better dataset for the models, this study focused on passenger car occupant injuries in two-vehicle car-truck rear-end crashes on divided highways. The well known five crash levels (KABC0) were employed in disaggregate models for 562 crashes involving 1,175 passenger car occupants. Two ordered probit models were developed to capture the qualitative differences between different crash severities. The first model only included main variable effects and no interactions. The second model included interactions among the main effect variables. It was found that the model with interaction effects such as cars being struck in the rear with high speed differentials performed better than the main effects counterpart. Crash event, driver, vehicular, roadway and environmental factors were found significant in these models.

Following the results of an earlier study, Mercier et al. (1999) examined the effect of age on injury severity, this time for broadside and angle collisions on rural highways. The sample selected consisted of 4,261 involvements in Iowa from 1986 to 1993. The study focused on driver-front passenger involvements on high-speed (55-65 mph) highways with different models for each point of impact relative to the vehicle occupants. Age was a significant factor for all of the models, while the use of lap and shoulder restraints was significant in reducing injury severity. The occupant position relative to the point of impact was not significant by itself, but as an interaction factor with age. However, in some of the models (front and back impact) relationships between the factors and injury severity were not significant for male occupants.

A different approach to the statistical modeling was undertaken by Chang and Mannering (1999) to find the relationships between occupancy and injury severity for truck and non-truck involved crashes. A total of 17,473 vehicle involvements from King County (including the city of Seattle) in Washington State were used in the analysis. In this case, the unit of observation was the most severe injury for each vehicle involved in a crash. The investigators developed a nested structure with four levels of vehicle occupancy in the upper nest and three levels of injury on the lower nest. This structure was used for modeling truck and non-truck crashes, resulting in eight models (upper nest). The results showed that most of the significant factors in the models had similar trends for both the truck and non-truck involved crashes. Non-use of driver restraints (safety equipment), ejection, and alcohol involvement were significant factors that increased injury severity. Other significant factors such as functional classification (non-truck) and speed limit (truck) were correlated, but showed the increased effect of speed in truck crashes. Gender was also relevant, with increased severity for females in truck crashes, but the opposite in the non-truck crashes. Young and old age groups were also found to have increased injury severity. Other significant factors included night time, season indicators and weekend indicators. Across the literature, weekend stands out as a consistently significant factor. This is probably related more to travel choice and driver behavior than to any significant change in environmental characteristics. The results also suggested that the truck involvement increased injury severity, more so on vehicles with multiple occupants.

In a study looking into a smaller and more homogeneous sample, Khattak (2001) used the 1994-1995 North Carolina data for rear-end crashes involving two and three vehicles and occurring on access-controlled highways. A total of 3,425 two-vehicle and 487 three-vehicle crashes were analyzed in three different models, one for each driver. For the two-vehicle crashes selected for analysis, driver 1 was the leading driver (no frontal vehicle damage) and driver 2 was the following driver (with frontal vehicular damage and no rear-end crashes). For the three-vehicle crashes selected, driver's 1 vehicle had rear-end damage, driver's 2 (following) vehicle had both frontal and rear-end damage and driver 3 (following) had only frontal end damage. This model selection sophistication was possible due to the limited nature of the data and information effects pursued in this study. In this case, three separate unrestricted models (one for each driver position) was preferable to using one (restricted) pooled model due to the impact of the sequence of drivers in a crash on injury severity. The results showed that the leading driver was more severely injured in two-vehicle crashes, while the driver in the middle is most severely injured in three-vehicle crashes. Being in a newer vehicle protects driver 2, while a newer vehicle 1 can reduce both driver's 2 and 3 injury severity. A few interactions among vehicle variables were tested and found to be non-significant.

In a regional study of the Province of Udine in Italy, Valent et al. (2002) developed several logistic regression models from a sample of 10,320 crashes involving 18,227 drivers in a six year period. The models featured adjusted odds ratios to assess relative risk of vehicle drivers (overall), car drivers, truck drivers, motorcycle riders, moped riders, cyclists and pedestrians. The odd ratio estimates for most models measured the likelihood of fatal injury vs.

non-fatal injury. For the car driver and truck driver models, the odd ratios were estimated for both the likelihood of fatal injury vs. no injury and non-fatal injury vs. no injury. The odds ratios for all drivers combined demonstrated decreased likelihood of fatal injury for females and increased chances for adult and older drivers (vs. younger drivers); drivers of motorcycles, trucks and bicycles (vs. car drivers); provincial and state roads (vs. municipal roads); and night hours (vs. day). Tendencies were similar for the motorcycle and moped riders, cyclists and pedestrians. Seasonal and day of week effects were also found to be significant, but changed trends by driver type, which shows some underlying travel choice exposure factors. For the car drivers some of the odd ratios (seat belt use and road type) tended to be much higher for the fatal vs. no injury models. All of the models were evaluated using the Wald test for the overall model and the Hosmer-Lemeshow test. The all driver combined model had the least favorable Hosmer-Lemeshow p-value, although the model had acceptable calibration. Another possible issue that may have influenced the higher odd ratio estimates was the long period (6 years) of data, which could cause overestimation due to repeated observations.

Using the crash data from the 1998 National Automotive Sampling System GES, Kockelman and Kweon (2002) developed six ordered probit regression models of driver injury severity. Three datasets were used: all crashes, single-vehicle crashes and two-vehicle crashes. Two models were developed for each dataset: all records (with and without speed variable) and a sub-set of records with the speed variable present. From the results, the age variable coefficient indicated increased injury severity with increased age for all crashes and two vehicle crashes, but not for single vehicle crashes. Although the continuous age variable does not provide full information about the effects for different age groups, these results suggest that the age effect tends to increase for the younger drivers and not for the older drivers. All the models agree on the coefficient signs for the crash types used in this study, with the head-on and rollover crashes having negative signs (increased severity) and the rear-end, left and right sideswipes with positive signs. The effect of vehicle types was different for the single and two-vehicle crashes. The results for the single vehicle crash model suggest that pickups and sport utility vehicles (SUV's) are less safe than passenger vehicles. On the other hand, pickups and SUV's have a positive (less injury) effect on their drivers and a negative effect on the drivers of the second vehicle involved. Other variables that were found to the significant included the presence of other occupants (increased severity for single vehicle crashes), daylight hour and interaction between weekend days and late-night. Since the GES sample is heterogeneous in geographical and infrastructural terms, the collection of variables and trends serves as a good lead for systematic studies of driver injury severity.

The results of studies in smaller jurisdictions, such as the one by Al-Ghamdi (2002), demonstrated a reduced number of significant factors. A sample of 560 subjects with serious (injury or fatal) involvements from August 1997 to November 1998 in the Riyadh, Saudi Arabia urban area were used for the model. A binary logit model of injury severity (fatal vs. non-fatal injury) was developed. Crash location (intersection vs. non-intersection) and contributing cause were found significant at the 0.05 level. However, the age variable coefficient had a p-value marginally significant (0.06). The crash location coefficient indicated that non-intersection involvements increased the risk of fatal crashes. This may be partially explained by the tendency of lower severity crashes at signalized urban intersections.

To remove possible confounding effects of seating position and vehicle characteristics, Bédard et al. (2002) used single vehicle crashes from the FARS database (1975-1998) to predict driver fatalities. The advantage of this data selection was the increased accuracy of the FARS database; however, it does not provide a generalized model for injury severity because there was at least one fatal injury in every crash and about 50% of the drivers sustained a fatal injury. A multivariate logistic model was developed and driver age, gender, blood alcohol level, point of impact, restraint use, the speed of vehicle, model year, and wheelbase indicator were found to be significant. This model correctly classified 69% of the observations. The seat-belt use (self-reported) may be biased due to legal consequences and when adjusting for a 14% over-reporting (based on previous studies), the odds of fatality were found to be 23% lower for drivers that used seat belts compared to those who did not.

Ulfarsson and Mannering (2004) developed several multinomial logit models to estimate contributing factors to injury severity. Four injury severity levels (no injury, possible injury, evident injury and disabling/fatal injury) were included in the multinomial structure. Fourteen separate models were used to estimate injury severity for male and female drivers in single vehicle crashes for passenger cars, pickups and SUV/minivans; two-vehicle crashes were modeled with the resulting four combinations of these vehicle types. For each model, as sub-sample was selected using exogenous, random sampling. A total of 22,068 driver involvements of Washington State crash data between January 1, 1993 and July 31, 1996 were used for the 14 models. The possible correlation in the two-vehicle accident cases because of two vehicles from the same crash was judged to be minimized by the sampling from a large database. The results showed that there are significant differences between the resulting models for male and female drivers in almost all of the seven cases. However, the vast majority of the coefficients retained the same direction (increasing or decreasing injury severity) in both the male and female models. Two notable exceptions are: contributing cause (did not grant right of way), and fixed object crash (struck guardrail). These are mostly related to individual driver behavior, thus major differences are expected.

An exploratory analysis of large truck crashes in California investigated the differences between rural and urban driver injury severity risks (Khorashadi et al., 2005). Crash records on 17,372 vehicles (11,072 involvements in urban areas and 6,300 in rural areas) were drawn randomly from the California 1997-2000 crashes involving at least one large truck. Measures were taken to avoid choosing more than one vehicle from each crash to avoid correlation problems. The multinomial logit model developed for the crashes in rural areas resulted in a wide variety of significant variables with 50 coefficients estimates. Overall model fit was assessed via the likelihood ratio index (ρ^2=0.52) suggesting a good fit. The most influential variables were crash location (beyond shoulder, left lane), vehicle movement prior to crash, alcohol involvement, highway location (intersection), vehicle type and large truck driver at-fault. On the other hand, the model developed for the crashes in urban areas resulted in a wide variety of significant variables with 55 coefficients estimates. Overall model fit was assessed via the likelihood ratio index (ρ^2=0.69) suggesting a good fit. The most influential variables were alcohol involvement, crash location (beyond shoulder), crash type (fixed object, broadside), driver at-fault (large truck and passenger vehicle) and vehicle age (older than 1981 model year). In addition to the strong likelihood statistical test results indicating that the two models should remain separate, there is compelling evidence of great differences between the two models. Not only the most influential variables and coefficients are very different, but a total of 13 coefficients were significant in the rural model, but not the urban and 17 coefficients were significant in the urban model, but not the rural.

A systematic methodology for a binary logit regression model was presented in a study of 73,746 pedestrian crash injuries in Hong Kong from 1991 to 2004 (Sze and Wong, 2007). The response was a fatal or severe injury vs. a slight injury. First, a main effects model was fitted to find the significant variables from a group of crash, pedestrian, environmental and vehicle characteristics available in the crash database. Then confounding and temporal interaction effects were explored to improve the model. The model goodness of fit was verified using the Hosmer-Lemeshow test and logistic regression graphical diagnosis, including leverage and residuals. The number of risk groups in the Hosmer-Lemeshow test was $g > p+1$, were p is the number of covariates in the regression model.

In general, injury severity analysis of involvements has been richly developed over the years to become one of the forefront tools used in the traffic safety field. The quality and accuracy of the models have improved as well as the interpretative power of these studies. In examining different types of studies the goal is to use them as a foundation to qualify and interpret the results of the analysis presented in this report. The limitations of crash report data are the greatest concern when studying injury severity, especially for minor crashes. The main analysis goal is to find the relationships between crashes without harmful interference from within crash effects. Some data subset modeling techniques included separate models by crash type, rural and urban, type of driver, gender, type of vehicle and driver age. In some of the studies, data preparation included separating single from multiple vehicle crashes. In others, driver involvements were selected by a systematic random process to avoid repeating involvement observations from one crash. In regards to injury severity levels, many of the studies reviewed in this section used combinations of the injury levels in order to maintain statistical power and avoid confounding effects.

Table 2-1: Summary of Injury Severity Regression Analysis using Crash Involvements

Published study	Statistical methodology	Correlation accounted*	Type of sample	Sample Size	Jurisdiction
O'Donnell and Connor (1996)	Ordered logit, probit models	No	Census	18,069	New South Wales, Australia
Mercier et al. (1997)	Logistic regression	No	Head-on crashes on high-speed roads	2,171	Iowa
Duncan et al. (1998)	Ordered probit	No	Two-vehicle car-truck rear-end crashes	1,175	North Carolina
Mercier et al. (1999)	Logistic regression	No	Broadside and angle, rural high-speed	4,261	Iowa
Chang and Mannering (1999)	Nested logit model	No	Truck and Non-truck vehicle crashes	17,473	Washington State
Krull et al. (2000)	Logistic regression	No	Single-vehicle crashes	59,743	Michigan, Illinois
Khattak (2001)	Ordered probit model	No	Multivehicle rear-end drivers (3 models)	3,912	North Carolina
Valent et al. (2002)	Logistic regression	No	Census	18,227	Udina, Italy
Al-Ghamdi (2002)	Logistic regression	No	Urban injury or fatal crashes	560	Riyadh, Saudi Arabia
Kockelman and Kweon (2002)	Ordered probit model	No	GES data (all crash types; 1,2 veh. crash)	N/A	National (U.S.A.)
Bedard et al. (2002)	Logistic regression	No	FARS data (single-vehicle fixed-object)	44598	National (U.S.A.)
Dissanayake and Lu (2002)	Logistic regression	No	Older driver fixed-object crashes	7,371	Florida
Toy and Hammitt (2003)	Logistic regression	No	Two-vehicle crashes	6,481	National (U.S.A.)
Ulfarsson and Mannering (2004)	Multinomial logit	No	Single and two light vehicle crashes	22,068	Washington State
Khorashadi et al. (2005)	Multinomial logit	No	Truck crashes (one driver per crash)	17,372	California

***Indicates whether statistical method accounted for correlation between involvements in the same crash.**

A summary of past studies where crash involvements were used in the injury severity analysis is presented in Table 2-1. The amount of past research demonstrates the capabilities of logistic regression as a method to model crash injury severity vs. a set of continuous and discrete independent variables. Interactions were found significant in several studies such as: light-weather, alcohol-seat belt, among others. Only a few models in past studies attempted to combine single and multiple vehicle crashes. No injury severity analyses were found to have exclusively used data from high-speed arterials. Although there is a substantial amount of literature demonstrating different uses of severity analysis, only a few studies dealt with a sample as large as the one undertaken in this research. Also, no past study has addressed the differences in reliability of the driver injury severity models when controlling for road entities and when controlling for road entities and crash types.

2.3 Crash and Injury Severity Model Comparisons

Several methods of assessing the validity of a model have been well discussed by the literature. This summary of previous works involving injury severity analysis is intended as a guide of the accepted methods to compare different statistical models. In general, logistic regression analysis can be assessed using several measures of performance. For comparing the reliability models that use different datasets, calibration and discrimination measures have been used. More on these measures is discussed in Chapter 3. Several studies have dealt with these comparisons, some of which are presented next.

In the study be O'Donnell and Connor (1996), the two models (ordered probit and logit) exhibited similar goodness of fit (Veal-Zimmerman Pesudo-R^2). The coefficients of the ordered probit were consistently lower than their logit counterparts. The asymptotic t-ratios suggested that the standard errors were lower for the ordered logit model, but this could not be verified. The coefficient signs of the two models agreed, except for the effects of time of crash, which were not found to be significant in the ordered logit model. In summary, none of the models showed a significant advantage over the other.

In a study using a sample of 43,913 crashes reported in Ontario, Canada during 1986 the investigators assessed the reliability of different crash severity models (Saccomanno et al., 1996). The criteria utilized included goodness of fit, robustness of risk factor coefficients, and whether the resulting coefficients where acceptable and consistent with previous research and scientific principles. In this study, three model structures were tested. Models 1a and 1b were disaggregating sequential binary logit models (five injury severity levels, four injury severity expressions). Two sequencing options (a, b) were developed: from No Injury to Fatal Injury, and vice versa. Model 2 was a disaggregate two-stage binary logit model, where two injury severities were considered at each stage. In stage 1, injury severities were classified as severe and non-severe. In stage 2, the severe cases were further classified into Fatal and Major Injury, whereas the non-severe cases were split into minor and minimal injuries and no injury. Model 3 was an aggregate binary logit model with only two severity levels: severe (fatal and major injury) and non-severe (minor, minimal, and no injury). The model comparison of statistical goodness of fit at the injury expression level employed a similarity index (to measure predictive reliability of each injury severity in each model) and the expected percent correct (case by case using Monte Carlo statistical estimating techniques). At the overall model level, two success index indicators that measured correct case-specific classifications in each severity model for each injury level as well as the whole model (all the injury expressions treated together) were developed. Also, the Predicted Less Observed Injury Severity Share was used for the overall model only. Each model was finally compared in terms of the statistical significance of the injury expression coefficients (t-tests), and whether the results were scientifically acceptable. The results of the analysis suggested that model reliability is not sensitive to the number of injury classes specified in the model or to the level of model aggregation. The most important factors explaining most of the variation in injury severity were the dynamics of the crash, seating position of the occupant, use of seat belts, and age of occupant involved. The accuracy of the information provided in the crash reports was the primary determinant of model reliability according to the authors.

Krull et al. (2000) used logit models to analyze injury severity for drivers involved in a single-vehicle crash. Three-year crash and road inventory data from Michigan (1994-1996, N=35,447) and Illinois (1993-1995, N=24,296) were collected from the Highway Safety Information System (HSIS) maintained by FHWA. The KABC0 injury scale is used in both jurisdictions and categories K (fatal) and A (incapacitating injury) were grouped together to represent severe injuries. Three single vehicle crash models were developed from the Illinois, Michigan and pooled data. A total of 16 driver, vehicle and environmental variables were included in the first regression analysis. However, *adt* was excluded due to a high correlation (0.533) with rural functional class, as shown by the correlation matrix. Likewise, the right shoulder width and left shoulder width were highly correlated and excluded as well. The significant factors found to increase injury severity for the three models were alcohol involvement, daylight, driver age, rural functional class, speed limit and rollover crash; on the other hand, restrain use, slick roadway and heavier vehicle types had a decreasing effect on severe injury. Only vehicle type showed non-significant coefficients for the Michigan and pooled models. When comparing the goodness of fit of these models, the pooled model performed better on the Likelihood Ratio Test statistic (not a formal goodness of fit measure), also the pooled model performed a little better on the R-square measure, which can be used to compare models using different datasets. Missing data-dummy variables for driver age and restraint use were found to be significant at the 0.10 level, suggesting a systematic reason for the missing variables.

A similar effort was undertaken as part of the final analysis presented in later chapters of this report.

In a study focused on driver characteristics (Dissanayake and Lu, 2002), two sets of sequential binary logistic regression models were developed to describe the injury severity relationship of older drivers involved in fixed object-passenger car crashes in Florida between 1994 and 1996. The dependent variable in one set of models was driver injury severity, while it was the crash injury severity for the other set. For each of the sets of models, crash or injury severity was varied from the least severe (no injury) to the most severe (fatality), and vice versa. The injury severity models were found to have better fit and predictive accuracy. The fit was compared using the rank correlation measures and the predictive accuracy was computed as the ratio of the true positives and true negatives to the total cases with a 0.5 cut point. The percent accuracy is equivalent to the percent concordant in a binary logit model.

A study by Abdel-Aty (2003) using data from three counties in Central Florida developed three driver injury severity models for different road entities. For roadway sections, crash data from 1996-1997 (17,647 drivers involved in 7,891 crashes) were used. For signalized intersections, the same crash data years were used with 2,336 drivers involved in 1,168 crashes. Meanwhile only the 1999-2000 police reports were available for toll plaza crashes for a total of 447 crashes and 803 involved drivers (725 with complete information). Different modeling methods were tested: multinomial logit, nested logit, and ordered probit. The goodness of fit measures likelihood ratio index and classification accuracy for each model were compared. The nested logit was the best model, while the ordered probit performed very well with considerable less data and modeling efforts. After testing several combinations, four categories of driver injury severity levels were found to produce the best models: no injury, possible injury, evident injury, and severe/fatal injury. The factors related to driver's age, gender, seat belt use, point of impact, speed ratio and vehicle type were found significant in all models. Driver at fault, land use and light-weather interaction were significant in the signalized intersection. Alcohol-seat belt interaction, lighting conditions, and the existence of a horizontal curve were found significant in the roadway section model. The model for toll plazas included weather condition, number of impacts, E-Pass lane, alcohol-seat belt, passenger car-speed ratio and two additional E-Pass interactions.

In a study of crash severity levels at signalized intersections, Abdel-Aty and Keller (2005) explored the differences between ordered probit models using complete datasets and restricted datasets. The complete data included short forms (minor crashes) and long forms (crashes available in the CAR and FDHSMV crash databases). The restricted dataset included only the long form crashes. Crash data from four counties in Central Florida during the years 2000-2001 were used to develop five models: 7,833 crashes reported on long forms (restricted dataset) and 21,204 crashes in the complete dataset (including short forms). The first two models analyzed the restricted and the complete dataset crash severity with only crash type and county indicator as independent variables. The next two models used the same datasets, but with intersection characteristics as their independent variables. In both cases, the models with complete datasets fared much better in classification accuracy. Also, right turn crashes were significant in the complete dataset model and not in the restricted model. Meanwhile, most gains in variable information were achieved in the intersection characteristics complete dataset model (major road no. of lanes, left and right turn lanes, division on minor road, and *adt* on major road). It was decided to use the complete dataset for the final model with a combination of independent variables of the previous models. This final model achieved a high level of classification accuracy (79.1%), but lost all but two (median and speed limit on minor road) intersection characteristic variables. The combined model also lost the right turn and sideswipe crash types (which are less severe). In this study, goodness of fit and variable information was used for comparison and demonstrated the usefulness of the complete dataset for the less severe crash types. It also showed a tradeoff between the amount of significant factors (especially road-related) and the overall risk assessment provided by the combined variable model.

A summary of the methods used to compare the statistical models just discussed is shown in Table 2-2. The coefficient signs are always checked for agreement with previous studies and scientific principles. Coefficient robustness and classification accuracy are also important in comparing models. This is not an exhaustive list, but the fundamental issues in comparing models have been adequately covered in this section.

Table 2-2: Summary of Goodness of fit Comparison Methods from Past Studies

Regression Analysis / Goodness of Fit	**O'Donnell and Connor (1996)**	**Saccoma-nno et al. (1996)**	**Krull et al. (2000)**	**Dissanaya-ke and Lu (2002)**	**Abdel-Aty (2003)**	**Abdel-Aty and Keller (2005)**
Statistical model Analysis type	Ordered probit and logit	Logistic regression	Multinomial logit	Logistic regression	Multinomial nested logit ord. probit	Ordered probit
Unit of analysis	Occupant involvement	Crash severity	Driver involvement	Driver involvement	Driver involvement	Crash Severity
Rank Correlation measures				X		
Classification accuracy		X		X	X	X
Likelihood Ratio Test			X			
Pesudo-R^2	X		X			
Coefficient signs	X	X	X	X	X	X
Coefficient robustness		X	X	X	X	X
Similarity Index		X				
Likelihood Ratio Index					X	

Finally, there are other studies that include additional goodness of fit statistics for the logistic regression model. Valenti et al. (2002) did not discuss model comparison, but provided the necessary goodness of fit data for the different logistic models. It was found that although all models have acceptable fit, the overall driver models had less favorable calibration (Hosmer-Lemeshow statistic) when compared to the models that only considered one group of drivers or pedestrians. The authors indicated that the study had several limitations, among them not having separate analyses for each crash type. In another logistic regression analysis (Sze and Wong, 2007) the model goodness of fit was verified using the Hosmer-Lemeshow test and logistic regression graphical diagnosis, including leverage and residuals.

2.4 Analysis of Crashes on Arterial Roads

2.4.1 Crash Frequency Analysis on Arterials

Several studies have focused on examining crashes on arterial roads and found important factors that influence both the frequency and severity of crashes occurring on these roads. Milton and Mannering (1998) used data for principal arterials in Washington State from 4,368 km. of highways and 11,757 crash records. Road characteristics by sections were integrated with the crash data. Sections with intersections were excluded due to lack of data about the crossing (minor) roads that could create omitted variable bias. Two negative binomial frequency models were developed for the eastern (11,058 sections) and western part (20,248 sections) of the state to account for climate variations. The observations, one crash count per road section, were entered into the models. The results showed that higher *aadt* per lane, Medium *aadt* (<2500), number of lanes, narrow right and left shoulders, tangent-curve indicator section increased the crash risk. Higher peak hour *aadt* percentage, truck (eastern) or single-truck (western) percent, speed limit, sharp horizontal curves, curve radius, and tangent length resulted in decreased crash risk. This shows that roads with higher exposure (in *aadt* and section lengths) higher conflicts (number of lanes) and lower design standards (narrow shoulders and long tangents with tight curves) increases the crash risks. On the other hand, congested roads, higher truck percentages, higher speed limits, curve radius and tangent length demonstrate higher design standards and better performance in crash frequency. Sharp horizontal curves tend to decrease crash frequency (perhaps due to driver caution), except when there are more curves relatively close (tangent-curve interaction).

Brown and Tarko (1999) developed a crash frequency analysis of total crashes, property damage (PDO), and injury/fatal crashes was developed using five years of data on 155 Indiana urban multilane arterial segments. The three negative binomial models agreed that the crash risk increased with increased access density (number of accesses per km of road) and the presence of signalized access points. The presence of outside shoulders, two-way left turn lanes (TWLTL) and medians without openings were found to decrease the risk of crashes in all three models. Further analysis showed that the percentage of injury and fatality crashes increased with the increase in access density in linear fashion. The effect of access control on crash severity was found weaker than on crash frequency.

Similar factors were found to be significant for the crash frequency model for one major arterial in Florida. Abdel-Aty and Radwan (2000) used roadway data from 566 segments which included 1606 crashes. A negative binomial model of crash frequency showed an increased risk with an increased section length, *aadt* per lane, degree of curve, and in an urban section.

Increased lane width per lane, shoulder and median widths decrease the risk of crashes. Additional models by driver age and gender showed additional factors that depend in driver involvement characteristics. For males, young and middle age drivers, an increase in the speed difference/speed limit increases the crash risk. Meanwhile, for older drivers a paved shoulder decreases the crash risk.

A study using road and crash data from two cities in Arizona focused on the median effects on urban arterial safety (Bonneson and McCoy, 1997). The negative binomial model results suggest that crashes are more frequent in segments with higher traffic volumes, driveway or street densities. Business or office land use increases crash risk when compared to residential areas. Regarding the median effects, the undivided cross section has significantly higher crash frequency than the TWLTL or raised curb median treatments when parallel parking was allowed in the undivided street. Meanwhile, where no parking is allowed in either street, the different between the undivided and the TWLTL treatments is small. The positive effect of raised curbs is greater for larger traffic volumes (>20,000 vpd). Research urban arterial roads in Ethiopia (Berhanu, 2004) showed that wider roadway width, wider and paved sidewalks and raised curb improves safety performance. Also, poor access management (increased minor junction density) negatively affects crash rates.

In a study using data similar to the focus of this investigation, Abdel-Aty and Wang (2006) focused on 476 signalized intersections along 41 arterial corridors in Florida. The results of this research showed significant correlations between nearby intersections. After dividing the intersections into 116 clusters, models to analyze the correlated data were developed. The results indicated that intersections with a large total number of lanes, high traffic volumes, short signal spacing, high speed limits along corridors, and a large number of phases are related with increased crash frequency. On the other hand, lower crash frequencies were associated with intersections having three legs, with exclusive right turn lanes on both roadways, having a protected phase for left turns in the corridor, and located in residential or open country areas. This study stressed the importance of the availability of geometric data for major and minor roads, traffic volumes and traffic control data for intersections. Signalized intersections on arterial roads are complex and additional information is required for the statistical analysis.

Potts et al. (2007) investigated the potential of right turn deceleration lanes in improving arterial operation and reducing crash risk. Economic analysis was performed using data from computer simulations with different major road volumes, right-turn volumes and right-turn speeds. A total of 602 scenarios with 30 simulation runs each were analyzed with costs for delays and crashes (based on Safety Analyst data) avoided with a 20 year life and a 4% minimum rate of return. The delays of right turns in four lane arterials are substantially lower than for two lane arterials, but the crash cost is higher for the four lane road. The main economical benefit of the right turn lanes for four lane urban arterial roads in this study was the safety improvement.

Arterial corridors have shown to pose some particular challenges when analyzing crash data. These types of roads have different types of access, higher traffic volumes and changes in land use which are characteristics different to other road types (i.e. two-lane or freeways). In addition, correlations between the nearby intersections and the lack of road data for some of the corridor components (intersections and driveways) are a crucial obstacle in obtaining accurate crash occurrence prediction. Traditional analysis by segments and intersections is not enough to describe the safety performance of these types of roadways.

2.4.2 Related Studies on Non-freeway Multilane Roads

Related studies with non-freeway and multilane roads suggest some of the factors that might be significant in crash occurrence and severity on arterial roads. Some studies have compared the percent reduction of crashes or crash rates under different road conditions. McCoy and Malone (1989) investigated the effects of left turn lanes on urban four lane roadways. The use of left turn lanes at intersections were found to significantly reduce rear-end, sideswipe and left turn crashes. On the other hand, on uncontrolled approaches of undivided roads, left turn lanes significantly increased right-angle crashes, while decreased rear-end, sideswipe, and left turn crashes. The authors argued that longer distance for cross-street maneuvers on undivided roads increased the right angle crash risk. This and other studies point to the importance of access management in reducing the most severe crashes in urban areas, by using medians effectively. In study of Indiana rural multilane roads crash rates Karlaftis and Golias (2002) found that, when controlling for *aadt*, median width and access control are the most important factors affecting crash rates followed by the pavement condition factors. Lane width and the presence of left turn lanes are also significant variables affecting crash rates.

A study of the crash rates and severities in 111 segments of high-speed (>40 mph) urban and suburban highways in Arkansas suggests that median treatments and access management has a positive effect on safety (Gattis et al, 2005). The analysis included comparing three years of crash data with road segment median, traffic volume and access frequency. Crash severity and types were also examined. As median widths increased, crash rates declined. As access density increased, so did the crash rates. Roads with shoulders and depressed medians had the lowest crash rates, while undivided roads with curbs had the highest rates. In terms of crash severity, the rates of severe (fatal and severe injury) crashes were significantly higher for undivided, without shoulder category, while adding shoulders significantly improved the crash rate. Eisele and Frawley (2005) before and after analysis had similar results. Crash rates and severities for 11 corridors in Texas and Oklahoma were compared. The crash rates increased as access density increased for all types of medians, while crash rates and severity decreased after raised medians were installed in the corridors.

In a crash frequency analysis of 10,517 segments of non-freeway state roads in Oregon a zero-inflated negative binomial model was developed (Strathman et al., 2001). The coefficients of the model showed a decrease in crash frequency when posted speed, average lane and right shoulder widths increase. The presence of left turn lanes, and vegetation or curbed medians also showed decreased crash frequency. Increased exposure (segment length), maximum curve length and maximum vertical grade increased the crash frequency in the model. Hauer et al. (2004) developed separate models for off-road and on-road crash frequency on urban four-lane undivided roads in Washington State. The fit depended mostly on variables *aadt*, number of commercial driveways, speed limit, while vertical alignment, lane and shoulder width have weaker contributions. The results suggests that horizontal curves of moderate degree are safer than tangent sections on urban four lane undivided road segments.

Finally, a recent study by Potts et al. (2007) using data from North Carolina, Minnesota and Michigan found no consistent, statistically significant between lane width and crash frequency. Separate negative binomial models were developed for urban and suburban arterials segments and intersection approaches for each state analyzed for a total of 180 models. Although the results were not consistent, in one of the models it was found that a lane width of 3.0 m (10 ft) or less in four-lane undivided arterial segments increased crash frequency. Meanwhile, lane widths of 2.7 m (9 ft) or less on four-lane divided arterial segments was found to increase crash frequency in another model. The lack of literature that addresses road characteristics in arterials does not allow for effective comparison between studies.

2.5 Injury Severity Involvement Contributing Factors

A series of driver, vehicle, environmental and crash factors have been analyzed in past research. Some of the most significant results related to the present work are shown in this section. Most of the injury severity analysis is undertaken with a narrower research focus compared to the crash frequency analyses in the literature. Due to its disaggregated nature, injury severity involvement is an appropriate unit of analysis to better understand the nature and mechanism of different crash types. A simple comparison of different studies focusing on certain crash types will bring useful insight for the analysis presented in this report.

A summary of some results from studies focusing on different crash types is presented in Table 2-3, page 43. The diversity of contributing factors presented here is not only a function of the differences in crash types, but is also dependent on other factors such as data availability and research focus. Some of the research studies focus primarily on driver characteristics, while others had roadway attribute data available. In spite of some of the limitations, this summary shows that driver characteristics are important for multiple crash types. Crash mechanism variables (including vehicle type) carry a major weight on the outcome of a crash. Roadway attributes have also an influence, especially on single-vehicle (most of the run off-road) crashes. Not indicated in this table, the different crash types studied by Mercier et al. (1997 and 1999) resulted in different coefficients carrying different weights. The information presented suggests that if we had few data limitations, significant differences between crash types can be expected and furthermore the models would give an insight into some of the contributing factors unique to certain crash types.

Table 2-3: Summary of Injury Severity Regression Analysis of Different Crash Types from Past Studies

Method / Type / Significant factors	Mercier et al. (1997)	Duncan et al. 1998	Mercier et al. (1999)	Lee and Mannering (2002)
Regression method	Logistic	Ordered probit	Logistic	Nested Logit
Crash Type	Head-On	Rear-end car-truck	Broadside, angle	Rural Run-off Road
Seating pos.	X		X	
Seat Belt / Safety Equip.	X	X	X	
Belt*Age	X		X	
Age	X		X	
Age*Position	X		X	
Position*Safety Equip	X		X	
Gender	X	X	X	
Vehicle type		X		
Speed Diff. / Speeding		X		X
Speed Limit		X		X
$aadt$/Lane		X		
Road Character		X		
Lighting Cond.		X		
Surface Cond.		X		X
Alcohol		X		X
Shoulder indicator				X
Narrow shoulder				X
Fixed-object				X
Vehicle Defect		X		
Rollover		X		
Time of day				X
Day of Week				X
Year				X

One of the early studies with a large and disaggregated cross section dataset (O'Donnell and Connor, 1996) resulted in a variety of road user, vehicle, environment and crash event contributing factors. Increases in the age of the occupant and vehicle speed led to small increases in the probabilities of injury and death. Other factors found to have an increased injury trend were: seating position (other than driver), blood alcohol level, age of vehicle, vehicle type (other than passenger car), vehicle make (one manufacturer), seat belt not worn, and type of collision (other than fixed object). The only surprising finding was the driver seating position being *safer*, which may underline seat belt usage effects. The author did not mention, but it was also interesting to see that the fixed object crashes were found to have lower injury odds ratio than the right angle crashes, which may be affected by vehicle occupancy (exposure) in single vehicle crashes.

The effect of light trucks and vans (LTV) on driver injury severity levels was investigated using a logistic regression model by Toy and Hammit (2003). A sample from the U.S. Crashworthiness Data System of 6,418 two-vehicle crashes was used in the analysis. The risk of severe or fatal injury (yes vs. no) was computed by means of the odds ratio. The Abbreviated Injury Severity (AIS) scale from 1 (minor injury) to 6 (maximum, untreatable injury); a value of 3 corresponds to *serious injury*. Vehicle factors for both involvements in each crash were first entered into the model. Driver factors were then added to the model to examined possible confounding effects, age and gender and restraint use did not appear to confound the driver injury severity relationships with vehicle factors. Crash factors were then considered and the delta-v (change in velocity, joint effects of vehicle mass and crash severity) acted as a confounder, especially for the injury risk of drivers involved in a pickup truck crash. The results of the final model showed increased risk when a driver is struck by a pickup, for older drivers, if no seat belt is used, for larger delta-v, for left side impacts. The risk of serious or fatal injury was decreased for rear impacts. The authors concluded that vehicle mass and crash severity contribute to the self-protection and risk to others in a crash. On the other hand, vehicle body type characteristics (stiffness and center of gravity) are underlying contributors to the injury severity risk.

Lefler and Gabler (2004) used national crash samples to investigate the fatality and injury risk of pedestrians stuck by different vehicles. Only accidents involving a single vehicle and pedestrians were included in the analysis. Data from the national FARS and GES databases, as well as the Pedestrian Crash Data Study (PCDS) were combined to increase the amount of information available for each crash. Pedestrians were found to have two to three times the likelihood of dying when struck by an LTV than when struck by a car. Pedestrian injury distributions showed a significant increase in serious head and thoracic injury when the striking vehicle is an LTV when compared to a passenger car, after controlling for impact speed.

In a study of 1999 Indiana crash data to analyze the effect of age and gender on driver injury severity, only single vehicle crashes were selected to avoid confounding effects of crashes involving multiple vehicles (Islam and Mannering, 2006). Three driver injury levels were used: fatality, injury and no injury. Six separate multinomial logit models were developed for males and females in each of three age groups: young (ages 16 to 24), middle-aged (ages 25 to 64), and older (ages 65 and older). The results confirmed previous studies that showed significant statistical differences between separate models by driver age and gender. Different contributing factors affect the injury risk for the gender and age groups modeled in this study.

The contributing factors for most models discussed in the previous sections are summarized in Table 2-4, page 46. The variety of the type of effects in the previous studies demonstrates the evolution of injury severity analysis into a formidable analysis tool. Interaction and confounder analysis is required especially when more factor groups are found significant in some models. Some of the effects of factors in multivariate analysis are subject to confounding effects by unobserved variables. The offset hypothesis predicts that driver's overconfidence (and riskier behavior) will negate some or all of the safety benefits of certain technological advances in vehicles. A study by Winston et al. (2006) tested the offset hypothesis using disaggregate data to analyze the effect of airbags and antilock brakes on crash outcome. The results of the analysis proved that airbags and antilock brakes had an insignificant effect on crash outcome probabilities. This suggests that many drivers trade off safety improvements for additional mobility.

Table 2-4: Summary of Significant Factors' Groups on Injury Severity Regression Analysis from Past Studies

Published study	**Factors found significant in the analysis**				
	Driver attributes	Vehicular characteristics	Roadway design attributes	Environmental factors	Crash characteristics
O'Donnell and Connor (1996)	Yes	Yes	–	–	Yes
Mercier et al. (1997)	Yes	Yes	–	–	–
Duncan et al. (1998)	Yes	Yes	Yes	Yes	Yes
Mercier et al. (1999)	Yes	–	–	–	Yes
Chang and Mannering (1999)	Yes	Yes	Yes	Yes	Yes
Krull et al. (2000)	Yes	Yes	Yes	–	Yes
Khattak (2001)	Yes	Yes	Yes	Yes	–
Valent et al. (2002)	Yes	Yes	Yes	Yes	Yes
Al-Ghamdi (2002)	–	Yes	Yes	Yes	Yes
Kockelman and Kweon (2002)	Yes	Yes	Yes	–	Yes
Bedard et al. (2002)	Yes	Yes	–	–	Yes
Dissanayake and Lu (2002)	Yes	–	Yes	Yes	–
Toy and Hammitt (2003)	Yes	Yes	–	–	Yes
Ulfarsson and Mannering (2004)	Yes	Yes	Yes	Yes	Yes
Khorashadi et al. (2005)	Yes	Yes	Yes	Yes	Yes

2.6 Literature Analysis Discussion

In conclusion, the disaggregate unit of analysis by driver involvements is a proven method to capture the most important contributing factors of injury severity of crashes. In the case of high-speed multilane arterials, there is a limited amount of literature that deals with the effects of driver-, vehicle-, roadway- and environment-related factors on crash frequency and severity. In addition, there are only a few studies that deal with a statewide sample of crashes focusing on a road type. Some studies such as Valent (2002) had several limitations because they did not perform separate analyses for each crash type and road entities. It was shown that different crash types result in different contributing factors affecting injury severity.

The literature shows the ample capabilities of logistic regression in driver injury severity analysis. In studies were the effects on two types of injury classes are pursed, it proved to be a powerful, yet flexible model that allowed analysis of injury severity vs. a set of continuous and categorical independent variables. Most literature mentions the use of one driver per crash and some such as Ulfarsson and Mannering (2004) pointed out the efforts to stratify the sample such that correlations between involvements will not bias the results. There is some recent literature that shows additional analysis of correlations between involvements in a crash. However, these are not mentioned in this discussion, as it is not part of the research goals in this phase of analysis, which is focused on the general method of analysis on high-speed multilane arterials. Future work should consider these additional relationships using the sampling methodology found best suited in this analysis.

None of the past studies have attempted to neither combine the single and multiple vehicle crashes in an injury severity model nor have compared the reliability of models defined by road entities and crash types, as proposed in the analysis in Chapters 4 and 5. The joint analysis considered in this investigation poses some specific challenges, including model stability, over dispersion due to clustering of crash types and possible bias different weights between the single and multiple vehicle crashes. These issues are addressed by comparing the joint analysis to the separate road entity and crash type analyses. Also, this study addresses a new paradigm in arterial crash analysis: a joint analysis of different road entities with traffic operations vastly different than on limited access or minor roads. There is recent evidence of spatially and temporal correlations that make signalized intersections at multilane arterials a challenging unit of analysis, as shown in studies such as the one by Abdel-Aty and Wang (2006). This may affect the results of joint analysis even when the selected statistical method in the present study does not directly accounts for spatial or temporal correlations. Current and future research on safer arterial corridors should be pursued to improve the method of analysis for the safety performance functions.

CHAPTER 3. RESEARCH DESIGN AND METHODOLOGY

3.1 Crash Data Description

3.1.1 Florida Traffic Crash Records Database

The Florida Traffic Crash Records Database is a compilation of crash report data maintained by the Office of Management Research and Development of the Florida Department of Highway Safety and Motor Vehicles. The information corresponding to the years 2002 to 2004 was obtained in a set of relational database tables in Microsoft Access® format. Three tables were used in this analysis: Events (crash information), Drivers and Vehicles. In these tables, driver, as well as road, environment and vehicle characteristics are provided for all crashes reported on long forms. The long form report is required by Florida law for the following cases:

- Motor vehicle crashes result in death or personal injury, or
- Motor vehicle crashes, in which a driver leaves the scene (hit and run), involve damage to an attended vehicle or property (Section 316.061 (1), F.S.), or
- Driving under the influence of alcohol, chemical substances, or controlled substances, as defined by the Florida statues (Section 316.061 (1), F.S.), or
- Other circumstances deemed important to the investigating officer; in particular when a vehicle is inoperable due to a crash (FDHSMV, 2007).

This constitutes a limitation in the number of reported crashes available in the database especially for property damage only collisions. However, some of this limitation can be overcome by selecting crashes on state roads only. State troopers are highly trained and tend to complete a large amount of long reports (about 40% of the total long report crashes). Since they operate on state roads, there is a higher probability of filling a long report for a crash occurring on a state road. However, the main reason for limiting the analysis to the state roads is to be able to use the traffic (*aadt*) and other road data available only for the state roads. This database is best suited for the driver information (involvements) because all driver data by crash report section are readily available. The long form crash report includes sections for each vehicle (or pedestrian) involved in a crash. Driver information is part of the vehicle section, thus it is also called a vehicle-driver section. Since our response variable is the driver injury severity, the Florida Crash Records Database was selected as the fundamental database for the data preparation.

The Florida Traffic Crash Records Database includes a site location field, which identifies a crash as being at an intersection, intersection-related and non-intersection. As defined in the crash database, intersection-related involvements are those occurring within a 250 ft radius of an intersection and that are related to the operation of such intersection. There are other classification, such as private property and parking lots, which are not considered in this investigation. The focus of the analysis presented here are the intersection (including the intersection-related), non-intersection and driveway crashes (site location codes 1-4) occurring at high-speed multilane (speed limit 40 mph or higher) arterial roadways. Severe crash driver involvements are defined as those with incapacitating or fatal injury (injury levels 4 and 5 in the crash report).

Previous research has demonstrated the relationship between driver, vehicle, crash and environmental elements and the risk associated with severe crashes given that a crash does occur.

A statewide analysis of high-speed multilane arterials involves a wide variety of possible risk conditions and thus the amount of variables that can be successfully incorporated in a model increases, given that enough sample size is available for each condition.

3.1.2 Crash Analysis Reporting System (CAR) Database

The FDOT Safety Office collects and maintains a crash database consisting of crashes reported on state roads. The data collection starts with the crash reports from FDHSMV which go through a filtering process. A data quality review of the reports may result in some changes as deemed necessary to the Safety Office. After this process, the crash data tables are merged with roadway characteristics tables created specifically for CAR. The road characteristics information source is the Roadway Characteristics Inventory (RCI) maintained by the Transportation Statistics Office. These tables are called *freeze-break tables* because of the process of obtaining the information. A snapshot of the RCI data is taken during the first days of January, applicable to the previous year. The program that forms the snapshot for the CAR System generates a handful of elements based on occurrence of point features in RCI along the length of the segment or break or based on a collective combination of other elements (FDOT, 2005). The roadway data available in CAR are more complete than that in the FDHSMV Traffic Crash Records database, but it is limited to crashes on state roadways. Another difference is that the crash report number is 9 digits long, adding one zero at the end of the FDHSMV report number. This was resolved by dividing the CAR report number by ten, making it equivalent to the FDHSMV number. For CAR crash reports with 8 digit numbers, additional verification confirmed that these did not match with 8 digits FDHSMV crash report numbers, rather these matched to 7 digit

FDHSMV crash report numbers. This validated the use of a divider to match crash report numbers.

The CAR database is stored on AS/400 mainframes, and accessibility is much more difficult when compared to its FDHSMV counterpart. Even though most of the crash report data are available, a total crash data download is almost impractical for statewide analysis encompassing multiple years of data. The total data extract would include multiple data table *breaks* in one text document, which have been found to be prone to errors, so additional programming and caution is required. In addition, many downloads will be required to acquire three years of statewide data. As an alternative, the CAR system has a menu of downloadable reports which include the most commonly used crash and roadway data. These reports include the augment detail extract (option 3) and vehicle driver passenger extract (option 4), which were the most complete reports available. However, an analysis of the information available for these reports vs. the FDHSMV Traffic Crash Records database indicated serious deficiencies. Crash information such as on-off roadway classification was missing. Driver's information was classified as at-fault and second driver and options 3 and 4 had to be merged to complete the information. Neither contributing cause nor harmful event information was available as required for this research. Therefore, a combination of data sources was used in this investigation.

Another report utilized in this investigation was the Intersection Reference Report. This report contains yearly crash rate information for all intersections on state roads by county and node number, which uniquely identifies an intersection on a state road in Florida. Separate reports can be downloaded for every year, with intersection information. One field of information was of interest: number of approaches (num_legs). This process will be discussed further in the data preparation section.

3.1.3 Roadway Characteristics Inventory (RCI) Database

According to the RCI Field Handbook (FDOT, 2005), the Roadway Characteristics Inventory is a database of various physical and administrative data related to the roadway networks that are either maintained by or are of special interest to the Florida Department of Transportation. This collection of highway information is maintained by District and the Transportation Statistics (TranStat) Office personnel and kept in the RCI database. The RCI elements that serve to define the roadway are comprised of components referred to as Features and Characteristics. These are the building blocks of the entire system, allowing access to the most complete roadway information in a series of reports in Microsoft Excel® format. This database has been web-enabled on the intranet platform of the FDOT.

The main use of this database was to filter the crashes that occurred on arterials. Features 121 (Functional Classification) and 122 (Road Access) were utilized. The characteristics used were funclass and rdaccess. A report limiting to state roads classified as arterials, but without full-access control was retrieved and used as one of the filters for the crash information. In addition, the *freeze-break* tables for the years 2002-2004 became available in time for the final analysis and roadway data were added to the dataset, see Section 5.4.1 for details.

3.1.4 Video Log Viewer Application

This application is also called the DOT Video Log and is maintained by the Transportation Statistics Office. It provides frontal and right views of any state road segment entered into the system. This ground level view can be framed backward and forward providing a continuous video or a particular high-quality digital snapshot of the road conditions. These videos can be searched by roadway id (the basic identifying element for any road segment) and milepost and road direction.

Figure 3-1: Video Log Snapshot of State Road 25 in Alachua County (RDWYID 26010000 North MP:3.577)

A photo snapshot of a state road intersection (median opening) is shown in Figure 3-1. This photo clearly shows traffic signs (speed, bike lane, and yield-one-way sign combination) and general roadway characteristics such as medians, shoulders and number of lanes. In a systematic statewide study sometimes there are small groups of roads with missing or incomplete data. The video log proved to be an important application during the data filtering process to distinguish between one-way and two-way roads. Also, it helped clarify some traffic control device issues raised during the course of the investigation.

3.2 Crash Data Preparation

3.2.1 Crash Data Preparation for the Exploratory Analysis

3.2.1.1 Crash Data Combination

As part of the preliminary analysis, data from one year (2004) were selected. The analysis performed was considered exploratory to examine the general trend in the crash data and train regression models. The process of merging the FDHSMV and CAR databases is presented in Figure 3-2, page 56. The driver and vehicle tables were merged (outer join) by crash report number and section number, which resulted in a driver-vehicle table with one row per involvement, with no loss of records. A simple merging operation (one-to-many) was performed to combine the events and the driver-vehicle tables in the FDHSMV data. This resulted in a combined dataset with one row per vehicle-driver involvement. Then, the additional road information from CAR was merged with the FDHSMV by crash number. No CAR database crashes were added in addition to the ones already in the dataset due to the driver and vehicle information limitations explained in previous sections.

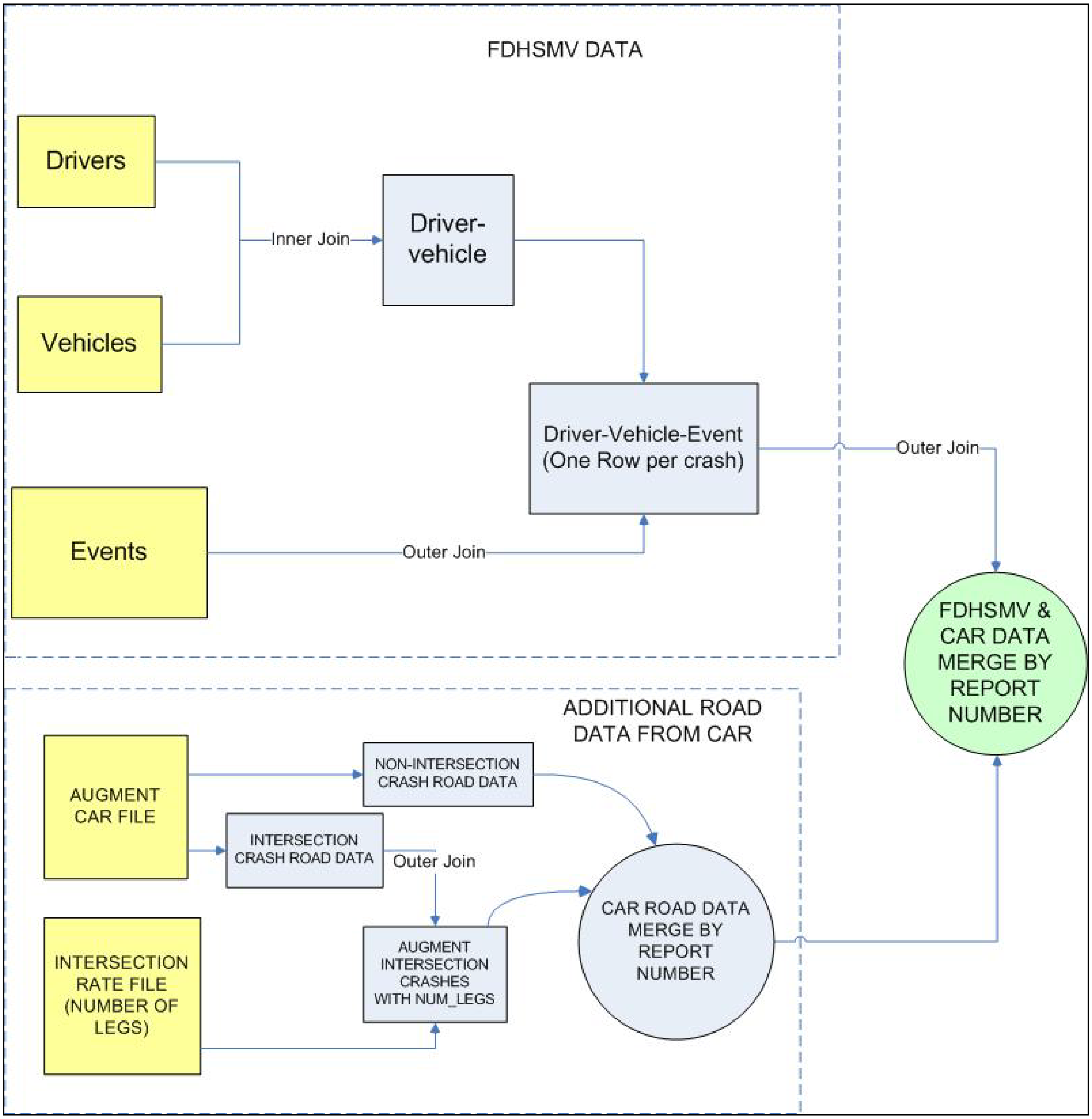


Figure 3-2: Data Merging for Exploratory Analysis (Year 2004 Only)

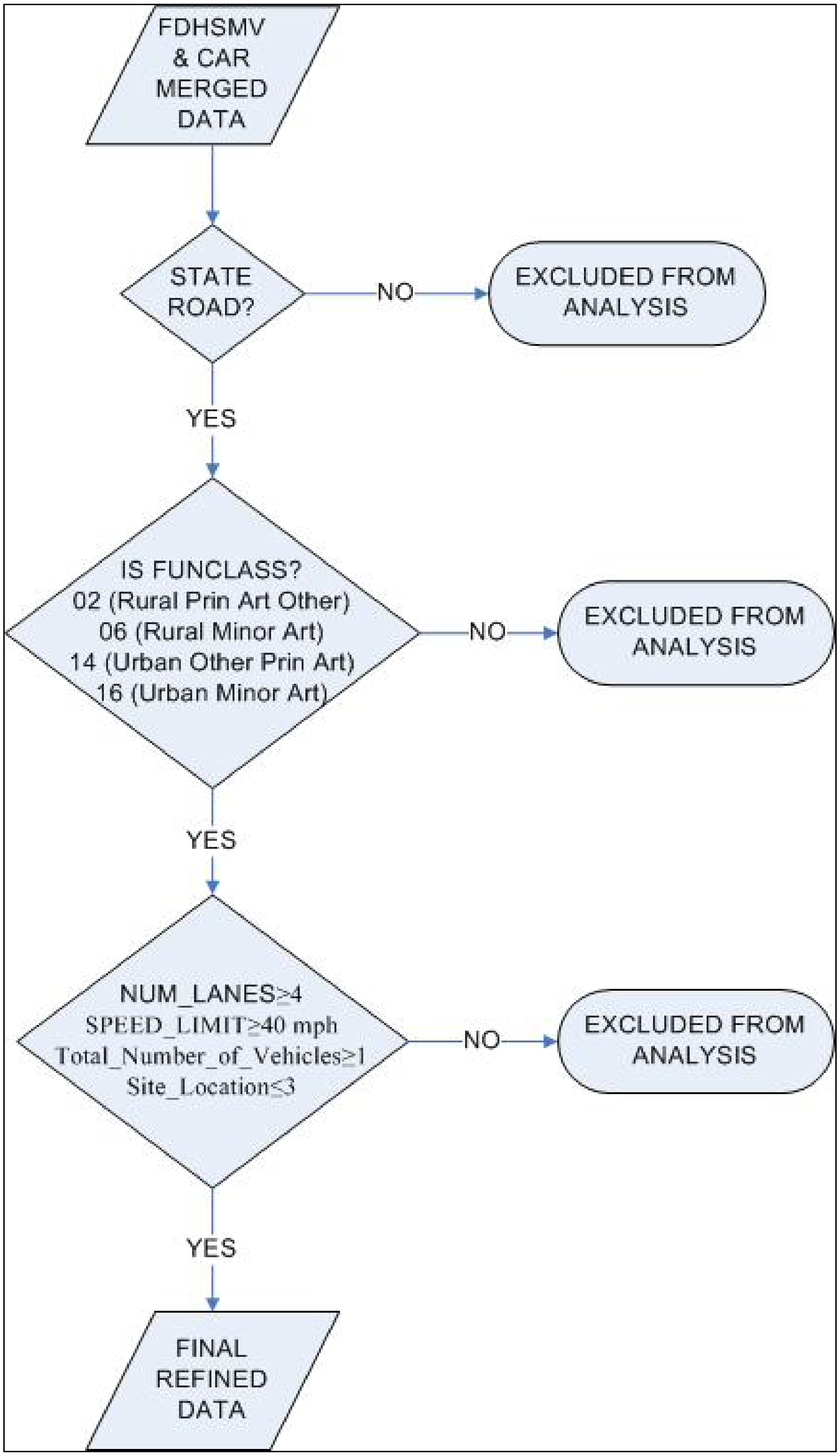


Figure 3-3: Data Filtering for Exploratory Analysis (Year 2004 Only)

3.2.1.2 Crash Data Filtering

There was a process of crash data filtering consisting of four steps. Figure 3-3, page 57, depicts the way the crashes occurring on high-speed multilane roads were selected. First, non-state road crashes were excluded from the analysis because of the scope of this research and the limited roadway attribute information available for these crashes. Then, crashes on non-arterial roads and freeways were excluded from the analysis using the funclass variable brought from the CAR database. Finally, the remaining arterial road crashes were filtered to include only the crash records with num_lane values equal or greater to four and speed_limit values greater or equal to 40 mph. Crashes occurring at intersections, influenced by intersections and non-intersection locations were selected for the analysis. Also, any remaining crashes were no vehicle was involved were excluded from the analysis by filtering total number of vehicles (>=1). This was only a precautionary measure since no driver involvements were expected in a crash not involving vehicles. The dataset used for the exploratory analysis consisted of 106,746 driver involved records. The final number of records used in the exploratory analysis was reduced to 60,221 records (involvements) due to missing data in both the FDHSMV and CAR databases. Missing data did not follow any pattern, and therefore it is safe to assume that the complete records subset was chosen at random.

3.2.1.3 Variable Coding

After the data were refined, each categorical variable was setup as part of the preliminary analysis, were the major cutoff points were determined by contingency tables, past research or data mining techniques (bin optimization for continuous variables). The categorical variables were coded using the reference cell method, were the level to compare against (base) was coded as the lower number (0 for a binary variable, 1 otherwise). Some of the continuous variables were later categorized according to the exploratory regression analysis results. The final analysis variable coding benefited from the experience in the exploratory analysis.

3.2.2 Crash Data Preparation for the Final Analysis

3.2.2.1 Purpose

After the exploratory analysis, it was found that very high-speed crashes may have influenced some of the results. Further investigation uncovered a situation in which the funclass variable value 14 (Urban Other Principal Arterial) included some limited access facilities, violating the high-speed multilane arterial definition put forth for this investigation. In response, a multi-leveled stringent filtering process was adopted. Also, due to the data setup, crash and roadway variables were repeated for multiple vehicle crashes. This would have a bias effect on some of the contributing factors. This would mostly affect variables that are significant to single vehicle crashes, some overestimation of the multiple (more than two vehicles) crash variables and correlation of unobserved effects. However, it was still effective in determining driver- and vehicle-related contributing factors. Since all involvements were considered in this analysis, it will serve for comparison with the final analysis.

A process for crash data preparation was outlined for the purpose of obtaining data from the two crash databases to be used in the analysis. It was later found that at least one of the crash databases did not contain enough information to correctly assign the crash records required for the analysis. Additional resources were also used in the process (see Figure 3-4, page 60). First, the two crash database data were merged into one larger database, excluding text and dummy fields that will not contribute to the analysis. Then, a filtering of crash records was applied to extract the high-speed multilane arterial crashes. RCI road lists and video log output was used in the filtering process and additional crosschecking assured that the filtering process was successful. Finally, additional coding allowed appropriate dummy variable and interactions to be entered into the model. Details about these data preparation processes are presented next.

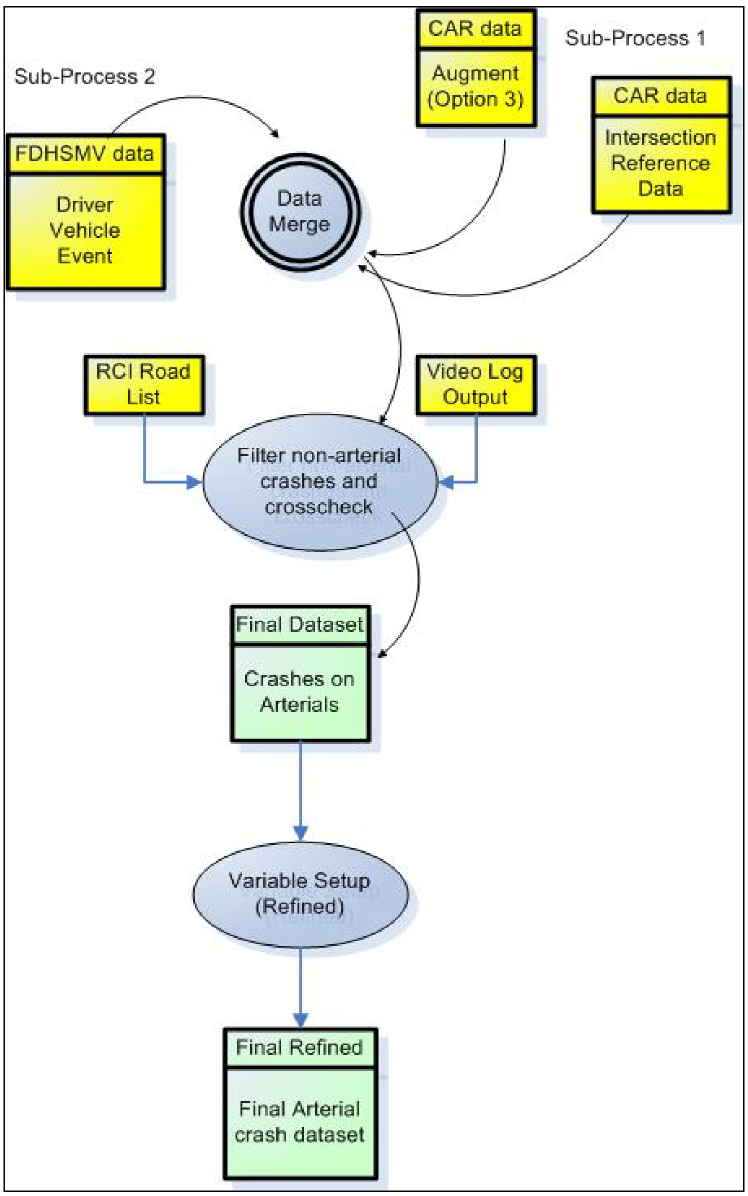


Figure 3-4: Data Preparation Process for Final Analysis

3.2.2.2 Crash Data Merge Process

3.2.2.2.1 FDHSMV Traffic Crash Reports Database Wide Formatting

Data preparation for the analysis included a series of steps to ensure that the most complete data available were used without introducing a bias in the sample. The first step was to accurately merge the tables in the access database provided by FDHSMV. The key index was the Accident Report ID, which linked the events table to the vehicles table. The vehicles table was in turn linked to the drivers table by both the report number and the section number (each section of the report has the information of each individual vehicle and driver involved in a crash). The vehicles and drivers data tables have one row per involvement. The drivers and vehicles tables were merged for up to four driver-vehicle involvements in each crash. Less than 5% of the crashes include more than 4 vehicles; these data becomes too sparse for analysis and thus were not included in the crash file. A graphical representation of this process is shown in Figure 3-5, page 62. After a wide format with up to four driver involvements per crash row is obtained, the driver-vehicle file is ready to merge with the events table by report number.

The major advantage of the wide formatting is that multiple cash involvements can be analyzed simultaneously. Due to the limitations of the analysis methodology and the scope of work of this study, the benefits of this format were not completely exploited. However, this data preparation process allowed a more complete analysis of variables by section number and the selection of an appropriate sample of crash involvements for the final analysis. The details of this analysis are presented in Section 5.3. The use of the wide format is recommended in future research work that includes severity analysis.

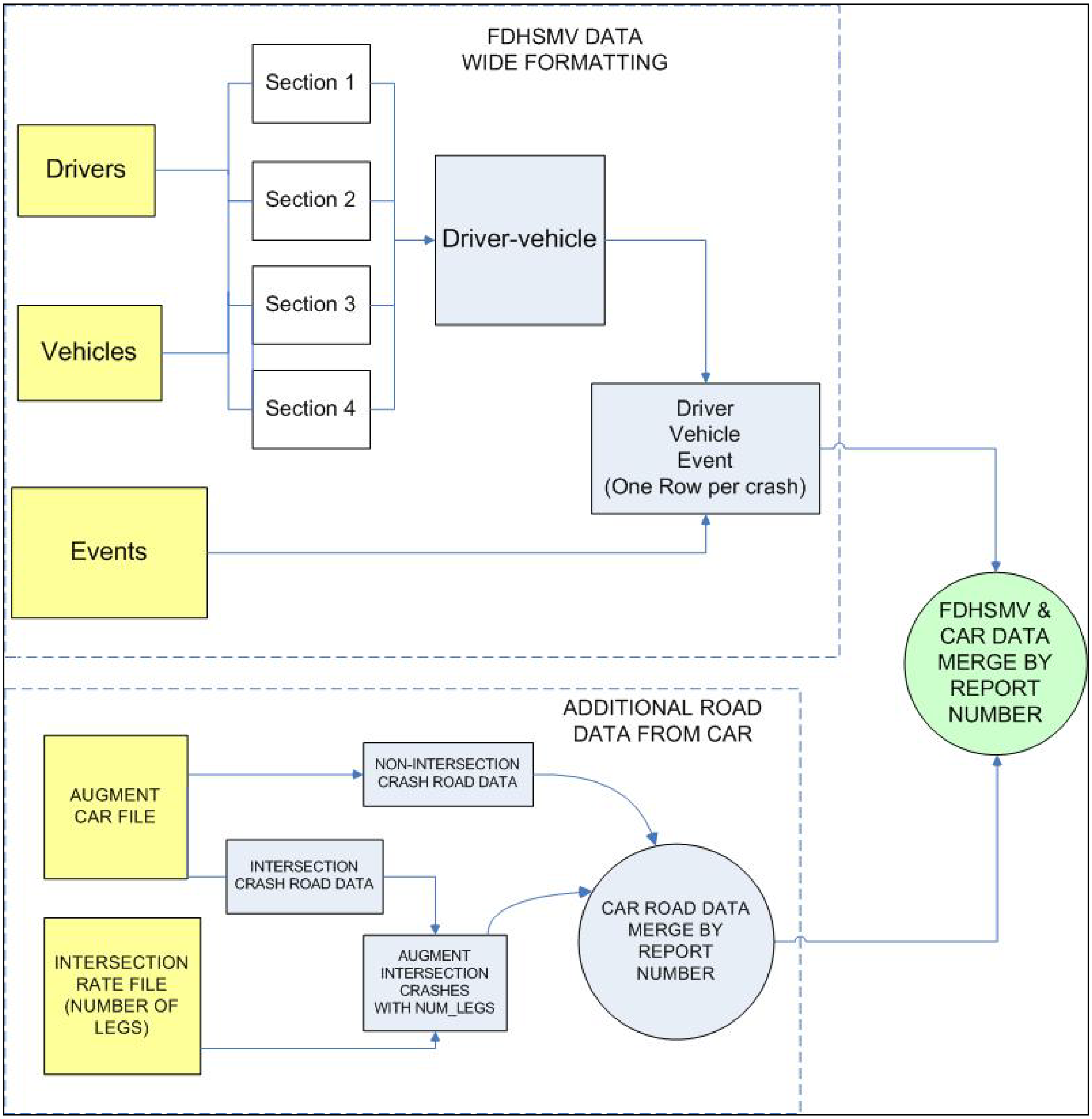


Figure 3-5: Process Chart for the Crash Data Merging

3.2.2.2.2 CAR System Additional Road Data

Once the FDHSMV crash data were merged successfully, an additional process was undertaken to add the traffic and additional road data from the CAR database. The augmented data file had additional geometric and traffic characteristics of the crash location. These include *aadt*, average truck volume percent, managing district, crash lane (as interpreted by the safety office, from the narrative/collision diagram), distance to the node (intersection center), Roadway surface (through lanes) width, first (closest to outside travel lane) shoulder width, Median width, and Skid Test Result number which are not available in the FDHSMV database.

The number of approaches (legs) has been found to affect the safety performance of intersections (Bauer and Harwood, 1996). In order to investigate the effects of this attribute, the information was first extracted from one report and then merged with the intersection crashes in the CAR augmented file. The previously described Intersection Reference Rates report was downloaded for each of the three years of analysis (2002-2004). A Microsoft Excel® macro was written to convert the downloaded text file format with breaks by page to an Excel worksheet with county, node number, and number of approaches. This file and the CAR augmented test file were imported into database tables in Microsoft Access® format. Then the number of approaches was added (for intersection crashes only) to the CAR augmented file by the county and node number via a make-table query in Microsoft Access®.

3.2.2.2.3 Final FDHSMV and CAR Crash Data Merge

The FDHSMV and CAR tables in Access® were imported into SAS® software datasets. These datasets were sorted by crash report number and then using a MERGE statement, the two datasets were combined into one dataset by crash report number. An outer join was used to include all FDHSMV records plus all CAR records that had a match with a FDHSMV record. Crash record repetition was avoided by comparing the records present in the individual datasets and the merged dataset. This merge resulted in the amount of 120,421 complete crash records for final analysis. This sample is more appropriate for analysis of single involvements per crash when compared with the 60,221 involvements for the year 2004 (which represents less than 30,000 crashes). In addition, there were approximately 28,437 CAR records not found in the FDHSMV data which amounted to approximately 13% of the total (28,437/215,898) records. However, these data did not include enough driver information for sections 1 and 2 and thus, not included in the final analysis. This discrepancy between the crash records in both databases is a concern that is being addressed by the Traffic Records Coordinating Committee of the FDOT.

3.2.2.3 Crash Data Filtering for High-speed Multilane Arterials

3.2.2.3.1 Crash Data Filtering Process

The crash record filtering process was effective in revealing some issues with the full access control road list and records were not imported correctly in Microsoft Access®. After these discoveries, this process (see Figure 3-6, page 66) was restarted to make sure all crashes were present. First, non-state road crashes were excluded. Then, the variable rdaccess was made available for the CAR crashes, therefore facilitating the process of excluding crashes on limited access roads. The crashes with rdaccess=1 (full-access control) were excluded.

Another check consisted of a subset of all crashes on one-way high-speed (≥40 mph) and multilane (≥4 lanes) roads to check if these crashes were indeed one-way. Since one-way roads tend to be in urban areas with lower speed limits, a small sample was expected. The Video Log outputs were used in this stage. Only 19 out of 436 (4.3%) records were two-way, the rest were one-way with less than four lanes on one side of the road, except for two cases in Miami. Due to the sparseness of these data and the different road conditions when compared to the high-speed multilane arterials, crashes on one-way roads were excluded from the analysis.

A list of freeway or limited access roads was obtained from a report in the RCI database. This list was then verified against the road1 and road2 fields in the FDHSMV in a partly automated fashion using several excel spreadsheets with crash report numbers. Next, groups of crashes in very high-speed (>65 mph) roads were examined to exclude those with a limited access facility. Many crashes (almost 4,000) not in the initial exclusion due to missing values in the road1 and road2 fields were captured this way.

Finally, the remaining arterial road crashes were filtered to include only the crash records with num_lane values equal or greater to 4 and speed_limit values greater or equal to 40 mph. The final analysis included crashes at intersections, influenced by intersections, non-intersections and driveways using the site location variable. Driveway-related crashes were included in the analysis as they are more common in urban arterial arterials and thus, an important part of this investigation. Also, crashes were no vehicle was involved were excluded from the analysis by filtering total number of vehicles (>=1). This was only a precautionary measure since no driver involvements were expected in a crash not involving vehicles.

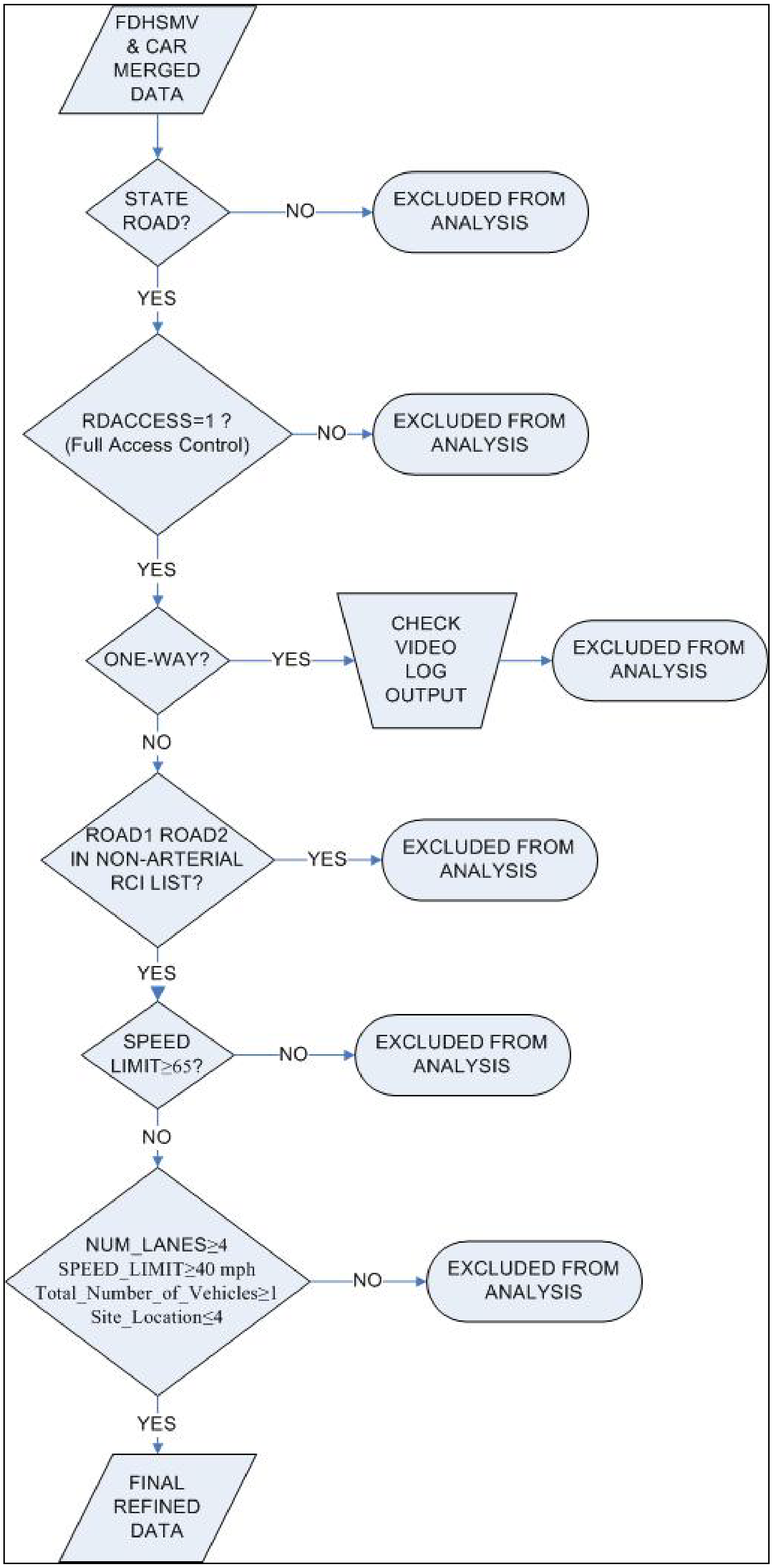


Figure 3-6: Process Chart for the Crash Data Filtering

3.2.2.3.2 Crash Data Filtering Crosscheck

After excluding the crashes on full-access control roads independent processes shown in the last three diamonds of Figure 3-6, page 66, served as crosschecks for crashes without the rdaccess variable (FDHSMV crashes that were not in the CAR system). The limited access facilities were excluded using the road1 and road2 variables, showing an excellent agreement with the rdaccess variable for the crashes with CAR records. Additional very high-speed (>65 mph) crashes were examined and only a small portion (93 out of 4,075 records) was excluded. This process allowed crosschecking of the previous filtering process. Although crashes without CAR records will not be included in the regression analysis, this initial filtering was useful in obtaining a second dataset with non-state high-speed multilane roads for future use.

3.2.3 Variable Setup

Before preliminary analysis, other variables interactions were derived from the original data. These included speed ratio, an indicator of speeding which consists of the ratio of the estimated speed divided by the speed limit. Another important interaction is the average traffic volume divided by the number of lanes. The *adt* per lane variable normalizes the volume to make different high-speed multilane arterials comparable. Previous research of signalized intersections on arterials using Generalized Estimating Equations (GEEs) statistical models for crash frequency analysis showed that the total average daily traffic per lane was the best representation of traffic volume (Wang et al., 2006).

Many of the crash report data fields that become variables in the analysis have a large amount of categories. Some are ordered according to the specific requirements of law enforcement in addition to those of the FDOT. Since the data may be sparse for many of the cells, after the categorical data analysis, some categories for dummy variables were combined. In addition, some variables suffered category rearrangements for them to have appropriate base categories for the injury severity analysis.

In addition, initially one-way contingency tables and other univariate descriptive statistics were used to assess the range of values in the database. Almost all of the variables had acceptable (valid) ranges. Extraneous values were investigated and removed as necessary to avoid outlier influence in the models. Overall, the quality of the data available was acceptable for the complete records.

3.3 Preliminary Analysis for Categorical Data

In order to find the basic relationships between groups of nominal and ordinal (categorical) variables, the contingency tables are utilized to summarize effectively the observed frequencies of each of the variables collected in crash reports and those derived from the report and other sources. The results of the simpler two-way contingency tables indicate the most significant associations (statistical dependence) between sets of individual variables. The variable of interest is driver injury severity. The contingency table analysis was performed using the SAS program PROC FREQ, which produced the tables of each variable with the statistical tests and measures of association to be discussed in the following section.

3.3.1 Contingency Table Independence Analysis

As part of the variable pre-screening process chi-square tests for independence were used to find the variables with statistically significant dependence to driver injury severity. This does not imply causality; it is a tool to pre-select a group of important driver, road, vehicle, and environmental variables to be included in the statistical model analysis. A two-way contingency table records counts of two characteristics, say X and Y, that can then be analyzed as shown in Table 3-1.

Table 3-1: Two-way Contingency Table Structure

	Y		
X	1	2	Total
1	n_{11}	n_{12}	$n_{1\cdot}$
2	n_{21}	n_{22}	$n_{2\cdot}$
Total	$n_{\cdot 1}$	$n_{\cdot 2}$	n

The Pearson chi-square statistic for two-way tables involves the differences between the observed and expected frequencies, where the expected frequencies are computed under the null hypothesis of independence. The chi-square statistic is computed as follows:

$$\chi^2 = \sum_i \sum_j \frac{(n_{ij} - e_{ij})^2}{e_{ij}} \tag{3.1}$$

where:

$e_{ij} = \frac{n_{i\cdot} n_{\cdot j}}{n}$ = expected cell values.

When the row and column variables are independent, χ^2 has an asymptotic chi-square distribution with (R-1)(C-1) degrees of freedom. For large values of χ^2, this test rejects the null hypothesis in favor of the alternative hypothesis of general association. In addition to the asymptotic test, PROC FREQ computes the exact chi-square test; however, this was not needed for the larger samples used in this investigation. Later, additional variables found to be important by previous research will be tested to find out possible confounder effects. In addition to the chi-square test of independence, the contingency coefficient is used as a tool to select variables with the strongest associations to driver injury severity.

3.3.2 Non-parametric Contingency Coefficient Measure of Association

Besides testing for independence, the contingency table analysis provides another useful tool to compare the relative strength of each variable relationship to the response, in this case the injury severity level. Measures of association serve the function of a correlation coefficient, some of which have the desirable properties of bound values (absolute: 0 to 1 or with sign: -1 to 1) and comparability (Kendall and Stuart, 1979). The Pearson's coefficient of contingency is perhaps the oldest measure used for non-numerical (nominal) variables. It was derived from the chi-square statistic of the independence test:

$$P = \sqrt{\frac{\chi^2}{n + \chi^2}} \tag{3.2}$$

The contingency coefficient has the advantage of normalizing by sample size to examine the degree of the association between two categorical variables. Since the contingency coefficient is based on the chi-square distribution, it varies by its degrees of freedom which are determined by the number or rows and columns in the table. For larger degrees of freedom, the chi-square

Probability Distribution Function is flatter and therefore larger values are expected for the same significance level. Therefore, the contingency coefficient does not have a bound and cannot be compared among different size tables. The attainable upper limit of P depends on the number of rows and columns. However, this measure can compare nominal and ordinal variables.

3.4 Preliminary Analysis for Continuous Variables

The first part of the preliminary analysis consisted of verifying the descriptive statistics using PROC UNIVARIATE in SAS. This included checking the crash reports for abnormally high of low values for the median width in the final analysis, which could be easily checked using the Google Earth™ satellite imagery application. This in conjunction with the data preparation cross-verification process was performed to ensure that records with invalid values were not entered into the analysis.

For variables with continuous distributions, correlation values were used to compare the relationships between different continuous variables and the driver injury severity variable. The Pearson product-moment correlation is a parametric measure of association for two variables. It measures both the strength and direction of a linear relationship, which is one of the characteristics of the final logistic regression analysis. The value of the sample Pearson product-moment correlation is defined as follows:

$$r_{xy} = \frac{\sum_i ((x_i - \overline{x})(y_i - \overline{y}))}{\sqrt{\sum_i (x_i - \overline{x})^2 \sum_i (y_i - \overline{y})^2}} \tag{3.3}$$

If one variable X is an exact linear function of another variable Y, a positive relationship exists if the correlation is 1 and a negative relationship exists if the correlation is -1. If there is no linear predictability between the two variables, the correlation is 0. If the two variables are normal with a correlation 0, the two variables are independent. However, correlation does not imply causality because, in some cases, an underlying causal relationship may not exist. The use of this correlation is limited to a few variables since most of the crash-related factors in the database are described qualitatively rather than quantitatively.

3.5 Regression Modeling

The use of binary logit (logistic) regression is well suited for crash and driver injury analysis, as shown in previous research discussed earlier. The binary response variable is the driver injury severity, which takes the values severe (response value=1) vs. non-severe (response value=0). Recall that the driver severe crash involvement is defined as those that resulted in an incapacitating or fatal injury. The logistic regression in this case models the probability of a driver sustaining severe injury given that he/she is involved in a reportable crash. The logistic regression model takes the form shown in Equation 3.4 as follows:

$$\text{logit}(p_i) = \log\left(\frac{p_i}{1-p_i}\right) = \alpha + \beta' X_i \qquad (3.4)$$

where:

p_i = $prob(y_i = y_i \mid X_i)$ is the response probability , and y_i is first ordered level of y;

α = intercept parameter;

β' = vector of coefficients to be estimated; and

X_i = vector of independent variables.

This is a linear additive model were the vector of independent variables are summed to obtain a total predicted probability of an event (success). The maximum likelihood estimators of the parameters yield the vector of coefficients. Additional details about the maximum likelihood estimation, the statistical significance tests and confidence intervals can be found in Hosmer and Lemeshow (2000).

3.5.1 Model Building and Assessment

The categorical variables were coded using the reference cell method, were the lower value was the base for the coefficient and odds ratio calculations. For the large amount of model fitting required for this investigation, the stepwise method for variable selection was selected. This is an acceptable method when properly used and permits fitting many models with fewer resources. For the model building, the stepwise variable selection available in PROC LOGISTIC was used to select the appropriate variables with 95% confidence levels limits for variable entry and removal. Once an appropriate main effects model was fitted, consideration was given to possible interactions. In all cases, PROC CORR was used to find the highest Pearson correlation factors among variables available in the crash involvement data. Various combinations of first order interactions were tested. The results were varied and for some models which resulted in poor goodness of fit (Homer-Lemeshow test p-value less than 0.05) or numerical problems appeared. The competing models were developed using the following steps:

1) Preliminary analysis of categorical and continuous variables.
2) Continuous variable categorization using data mining bin optimization techniques.
3) Stepwise method for logistic regression main effects selection.

4) Check for Akaike Information Criterion (AIC) differences in the intermediate steps detail for each model.

5) Use Pearson/ Spearman correlations and scientific inference to test for first order interactions.

6) Additional stepwise regression for interaction selection (where necessary).

For better model fitting results, the detail option of the stepwise method was used to examine the AIC for each intermediate step, to have a secondary comparison method similar to the best subsets method. For the larger models, some variables that fulfilled the 95% significance levels of the stepwise method, did not have justifiable (>10) differences in the AIC statistic. In such large models this resulted in some over fitting, therefore the stepwise variable selection limits were more stringent for the larger sample models. It was decided to use 98% significance levels for the six road entity and traffic control models. After all the models were developed, the model interpretation, assessment and comparison processes were performed as detailed in the following sections.

3.5.2 Interpretation of Coefficients using Odds Ratios

One of main advantages of the logistic regression model is the ease of interpretation of the estimated parameters. This is achieved by calculating the odds ratio for each of the coefficients in the model. This statistic represents the probability of certain outcome given a fixed effect value, relative to a base value. The estimated odds ratio is obtained by exponentiation of the logit difference, as shown in Equation 3.5:

$$\hat{O}R(a,b) = \exp[\hat{\beta}_1 \times (a-b)] \qquad (3.5)$$

If the variable is binary and coded (0, 1), this reduces to $\hat{O}R = \exp[\hat{\beta}_1]$, the odds ratio estimates and their 95% confidence intervals are computed by the SAS software (SAS, 2007).

The interpretation of this variable is straightforward. For continuous variable, the odds ratio represents the change in probability of certain outcome by one unit increase of the effect, given that every other parameter is fixed. For categorical variables, it represents the probability of certain outcome when the status (category) of the effect changes relative to the base value, given that every other parameter is fixed.

3.5.3 Model Assessment

After all of the competing models were developed, the models were investigated in terms of both the statistical fit and the resulting coefficients. The goal was to obtain a model (or set of models) that strikes a balance between fit and simplicity (parsimonious principle). To achieve this, we must guard against either over fitting or losing important crash involvement factor information in the sake of simplicity. Several criteria have been developed to help choose a model or set of models, some particularly suited for logistic regression of binary data. The Akaike Information Criterion (AIC) is used to identify the best approximating model among a class of competing models with different numbers of parameters. The AIC has the form shown in Equation 3.6:

$$AIC = -2 * ML + 2 * k \tag{3.6}$$

where:

ML = maximum likelihood (log-likelihood); and

k = the number of variables in the model.

This parameter has the desirable property of being an efficient model selection criterion for large samples. In other words, the errors made by predicting the response variable using the model selected by AIC approaches the error of the best possible theoretical model asymptotically (as the sample gets larger). In general, if values of the AIC between two models differ by more than 10, the model with larger AIC has considerably poorer fit, and would normally not be considered further (Simonoff, 2003).

In addition, the performance of the model itself must be assessed. To achieve this in the case of the logistic regression model, a closer look at both calibration and discrimination is warranted. Fortunately, the SAS® software used for the multivariate binary logit modeling has various summary measures of goodness of fit. A popular overall goodness of fit is the Hosmer-Lemeshow statistic. This test measures the calibration of the model by grouping the observed and expected frequencies within each decile of risk. Then the resulting contingency table is tested for the chi-square function of variable independence to prove that the model has a good fit. In addition the Receiver Operating Characteristics (ROC) curve has valuable information for discrimination purposes. In addition, the areas under this curve of 1 minus Specificity (x-axis) vs. Sensitivity (y-axis) are used to determine how good the probability of correct classification is. Both of these measures plus the other traditional measures of fit better describe the model's performance (Hosmer and Lemeshow, 2000).

The estimated covariance matrix of the logistic model can point to possible multicolinearity problems in the data. These were examined and as a result some problematic variables were eliminated, with significant improvements in the models goodness of fit.

3.5.4 Relative Variable Significance

The type III analysis available as part of the PROC LOGISTIC output tests the significance of the effect of each factor *added* to the model containing *all other factors;* that is, to find the proportional or *relative* contribution of the factor to the explanation of the dependent variable (Le, 1998). The type III chi-square value for a particular variable is the difference between the generalized score statistic for the model with all the variables included and the generalized score statistic for the model with this variable excluded. The hypothesis tested in this case is the significance of this variable given that all the other variables are in the model. The small p-value indicates that the effect of this variable is highly significant (SAS, 2007). By utilizing the Type III test, we can compare the most important variables found in each statistical model to derive relationships between the factors affecting injury severity and the road, traffic, driver, and possible environmental differences found at each of the road entities represented by each statistical model.

3.6 Driver Involvement Modeling by Road Entities and Crash Types

The model series is divided into two recognized crash analysis schemes: by road entity and by crash type. The first requires analysis of crash data using different combinations of road entities; i.e. location (segment, intersection) and traffic control (signalized, unsignalized). To better recognize the different crash mechanics of involvements occurrence at different locations under different traffic control, the analysis tree shown in Figure 3-7, page 78, depicts the six different model schemes tested by using statistical modeling of severe vs. non-severe driver injury crash involvements. The six models are: all involvements, all intersections, signalized intersections, unsignalized intersections, pure segment, and segments (pure segments plus unsignalized intersections). The relationships between the data subsets used in the first six models are represented by the lines connecting the leaf elements. The broken lines represent that the unsignalized intersections were tested both as part of the intersections and segment models and as a separate model. By testing the reliability of each model and comparing it to the others, additional knowledge was acquired about the best way to model crash data for the high-speed multilane arterials. In addition, the comparison of the information about the contributing factors for each model will help recommend measures directed to reducing the severe crashes.

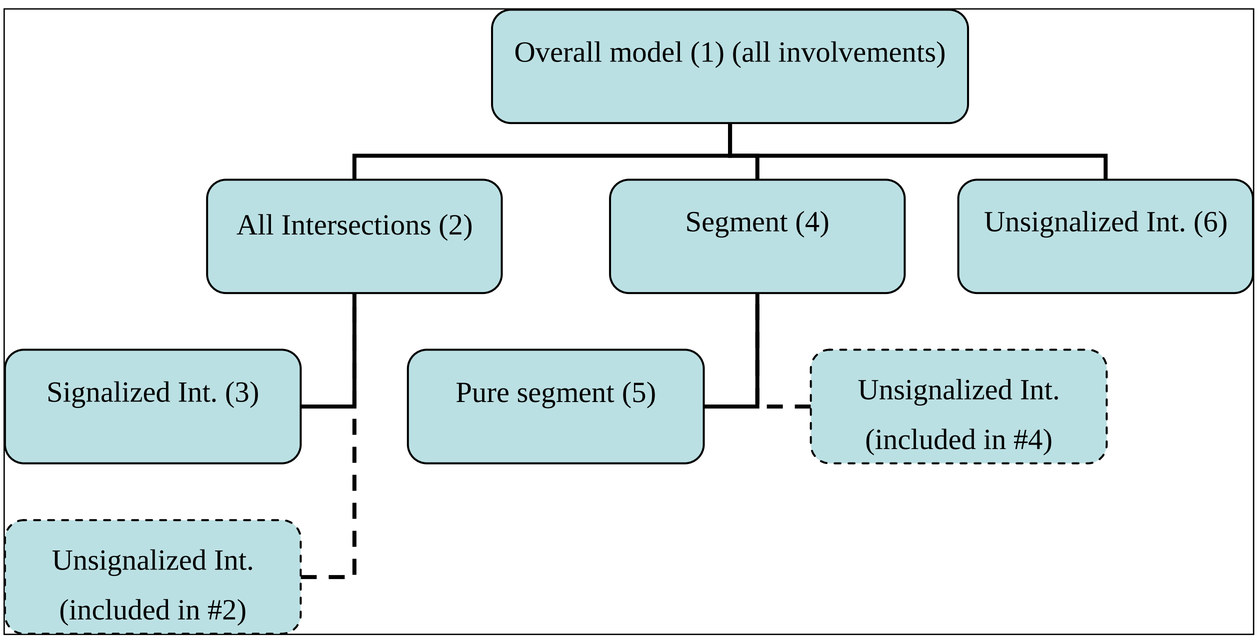


Figure 3-7: Driver Injury Severity Analysis Tree by Road Entity

The second severity crash analysis scheme uses different samples by crash type combined with the road entities to confirm and enhance the results of the first analysis. The crash collision type strata are shown in Table 3-2, page 79. For each specific crash type and using the previously designed road entities, a more accurate subset of crash involvements on high-speed multilane arterials can be investigated. This is expected to provide additional information not found in the first analysis by road entity. Also, possible correlation effects between crash types and driver injury severity can be addressed at this point. Analyzing crash driver involvements using a dual strata classification is intended to provide a more clear picture of the contributing factors in terms of driver behavior and vehicle characteristics as well as other environmental and road characteristics that affect the likelihood of a crash occurring and the level of injury as a result of such event.

Table 3-2: Driver Injury Severity Analysis Matrix by Crash Type and Road Entity

Entities/ Collision type	Signalized	Unsignalized	Segments plus Unsignalized	Purely segment	Overall Model
Rear-End	Model 1	Model 2	Model 3	Model 4	Model 5
Angle	Model 6	Model 7	Model 8	Model 9	Model 10
Left Turn	Model 11	Model 12	Model 13	Model 14	Model 15
Fixed Object	Model 16	Model 17	Model 18	Model 19	Model 20

3.7 Regression Goodness of Fit Comparison

The model comparison is very important to the objectives of this research. A great deal of effort was invested in data preparation and model fitting such that the different models would have equivalent variables and variable selection process that could be compared. Although individual models were flexible, the total variables tested for all of the models were the same. More than other measure of goodness of fit was used for model comparison. In addition, other factors were considered when comparing the models.

As explained in Chapter 2, an acceptable driver injury severity model should comply with the following principles:

- The models should have acceptable goodness of fit.
- The coefficients must be robust (statistically significant) to an acceptable p-value.
- The coefficient values (signs) for the contributing factors should agree with scientific principles and previous research results.

These principles guided the process of model comparison in this investigation. The goodness of fit measures for calibration (Hosmer-Lemeshow) and discrimination (ROC curve) were used for comparison. In addition, the AIC statistic and the R^2 value were considered. The coefficient robustness (p-value) was made more stringent for the models with larger sample sizes to compensate for this advantage (and avoid over fit) in order to have comparable results. Both the quantity and the relative importance of the coefficients for each model were examined in the comparison. Finally, the coefficient values (and standard errors) were examined in light of the scientific principles and previous research results.

CHAPTER 4. EXPLORATORY ANALYSIS

4.1 Significant Factors of Categorical Data Analysis

The first part of the analysis used crash, road vehicle, driver and environmental characteristics for each driver involved in a crash. The purpose of this is to find the major factors associated with increased injury caused by a crash. This analysis attempts to explain the likelihood of severe injury. Part of this analysis of each crash occurrence will focus primarily on the crash, road and environment characteristics.

The results of the preliminary two-way analysis of all of the crash involvements indicated rejection of the null hypothesis of statistical independent variables (against driver injury severity) for all of the initial 33 variables of interest. Among the variables associated with driver injury severity, first safety equipment (seat belt use), ejected and first harmful event (type of collision) had a significant effect. Other variables hold relatively good associations, although no direct comparison may be possible due to differences in the degrees of freedom, as explained previously. In addition there are some variables that are not as strongly associated, but may become important confounding effects.

Additional analysis of variable interactions is possible by using three-way contingency tables, where there is a ‘control’ variable. This allows close examination of the trends of certain events or characteristics (such as driver injury severity) with other variables. However, the analysis cannot go any further into variable interaction. Even with large sample sizes, at this point some expected cell values become too low to have interpretative value. On the other hand, statistical regression models provide more complete interpretative and prediction power, although with an error margin. The contingency table analysis is therefore the basis to find significant statistical relationships as well as a tool to interpret specific variable behavior in a sample.

There is fundamental difference in crash involvements between intersections and road segments. To investigate the validity of this claim with the data under study, three-way contingency tables were analyzed using the Cochran Mantel Haenszel (CMH) general association statistic. Each set of tables control the site location type (intersection or non-intersection crashes) for each of the 33 driver, road, vehicle, crash and environmental variables against the injury severity level. The results of this analysis, to be shown in both Table 4-2, page 85, and Table 4-10, page 85, indicate a strong general association (p-value <0.001) for 29 variables in the non-intersection and 25 for the intersection and intersection-related involvements.

The preliminary two-way statistical analysis was divided into separate groups of intersection and non-intersection crashes. This analysis includes crashes occurring on Florida's state roads. A total of 147,866 involvements occurred during the year 2004. This total was filtered down to 106,746 involvements not on interstates, freeways or expressways and only those identified by the Site Location ID as: Not at intersection, Intersection or Influenced by intersection. Those involvements occurring at intersections or related to intersections numbered 64,972 and those occurring at segment sections of the road amounted to 41,774. Crashes occurring at other types of site location were not included in this analysis. Further analysis on the individual contingency tables for intersection and non-intersection crashes will be stratified to find important relationships changes for the in-depth analysis stages. A summary of the total driver's involvements by injury severity level is shown in Table 4-1, page 83. This total was later reduced to 60,221 crash involvements in high-speed (≥45 mph) multilane (≥4 lanes) roads for use in the exploratory regression analysis.

Table 4-1: Groups of Driver Crash Involvements on State Roads in Florida during 2004

Site Location ID	Driver Injury Level					Total	Percent
	PDO	Possible Injury	Evident Injury	Incapacitating Injury	Fatal Injury		
Non-Intersection	26433	7688	5078	2180	395	41774	39.13%
Intersection (all types of traffic control)	29509	10692	7850	3050	250	51351	48.11%
Influenced by Intersection	9531	2297	1337	417	39	13621	12.76%
Total	65473	20677	14265	5647	684	**106746**	100.00%
Percent	**61.34%**	**19.37%**	**13.36%**	**5.29%**	**0.64%**		

The statistics for the state roads in Table 4-1 show that 57.7% of the fatal injuries occur on road segments, while 61.4% of the intersection (or intersection-related) involvements resulted in incapacitating injuries. These results suggest that the severe crash group brings a more complete picture, showing that the majority of the severe injuries result from involvements at or related to intersections. In general, intersection and related to intersection crashes are the majority (61.9%) of 2004 driver involvements. This underscores the importance of examining the intersection crashes. However, since this analysis deals with arterial corridors in a systematic fashion the relationship between intersection crashes and the surrounding environment is very important. A comparison between the contributing factors between intersection (including intersection-related) and non-intersection crashes is a first step to shape the rest of the analysis.

In addition to site location types, crash contributing factors usually change according to the type of road in which it occurs. General association CMH statistics were tested for the candidate variables vs. injury severity when controlling for Urban ID, Lane Group, Speed Limit Group and Traffic way Character (straight or curve section). The results of this analysis indicate strong general association (p-value<0.005) across all the variables, except for month, similar to the previous results. When different types of roads are examined, preliminary analysis suggested a difference among the relationships related to driver involvements and injury severity.

4.1.1 Analysis of Non-intersection Crashes

4.1.1.1 Non-intersection Cashes in All Non-limited Access Roads

The group of non-intersection crashes was organized in contingency tables to test whether each variable is statistically independent. The variables that showed dependence were then ordered by minimum number of rows or columns in their corresponding tables to make possible some comparison of their relative strengths of association (see Table 4-2, page 85). Each of the variables listed was cross tabulated against the five level injury severity variable previously explained in Section 2.1.2.

Most parameters showed statistically significant dependence with 29 out of the 33 variables tested. These variables can then be used for the non-intersection model fitting detailed in the next sections. Variables that appear shaded have stronger associations (measured by contingency coefficients) among similar size tables. Additional variables with weaker associations, yet with significant statistical dependence are also considered. These variables are road and driver characteristics that may be found to have interaction effects, as shown in the stratification analysis (see Table 4-7, page 90).

Table 4-2: Preliminary Analysis for Non-intersection Crash Driver Involvements (N=41,774)

Variable	Contingency Coefficient	DOF	Min (r,c)	Cramer's V	chi-square	p-value
Ejected	0.323	4	2	0.3413	4866.54	<.0001
Off/On Roadway	0.1965	4	2	0.2004	1678.019	<.0001
Speeding ID	0.1445	4	2	0.146	890.6299	<.0001
Urban ID	0.1298	4	2	0.1309	715.5656	<.0001
Alcohol Drugs	0.1239	4	2	0.1249	651.8123	<.0001
Traffic way Character	0.1087	4	2	0.1094	499.6669	<.0001
Vehicle Fault Code	0.1047	4	2	0.1052	462.5921	<.0001
Gender	0.0934	4	2	0.0938	367.6913	<.0001
Lighting Condition	0.0654	4	2	0.0654	178.4041	<.0001
Undivided Highway	0.0504	4	2	0.0505	106.4763	<.0001
Bad Weather	0.038	4	2	0.038	60.4578	<.0001
Crash Damage (Vehicle)	0.3139	8	3	0.2337	4564.618	<.0001
First Safety Equipment	0.3025	8	3	0.2245	4209.063	<.0001
Vehicle Type (41,225)	0.2925	8	3	0.2163	3856.386	<.0001
Location Type	0.2062	8	3	0.149	1854.72	<.0001
Speed Limit Group	0.1697	8	3	0.1218	1239.117	<.0001
Driver Action	0.1161	8	3	0.0827	571.0572	<.0001
Type of Shoulder	0.1108	8	3	0.0789	519.5882	<.0001
Road Surface Condition	0.0195	8	3	0.0138	15.8208	<.0001
Lane Groups	0.1142	12	4	0.0664	551.828	<.0001
First Contributing Cause	0.0841	12	4	0.0488	297.8924	<.0001
Time Group	0.0779	12	4	0.0451	255.3446	<.0001
Race	0.0602	12	4	0.0348	152.0791	<.0001
Location on Roadway	0.2064	16	5	0.1055	1859.057	<.0001
First Harmful Event	0.1682	16	5	0.0853	1216.571	<.0001
Number of Lanes	0.1401	20	5	0.0707	836.3189	<.0001
Vehicle Movement	0.116	32	5	0.0584	570.2234	<.0001
Day of Week	0.0542	24	5	0.0272	123.2607	<.0001
Driver Age Group	0.0444	28	5	0.0222	82.5573	<.0001

Examples of cross tabulation tables used in the categorical data analysis are presented next. Some of the variables with the highest chi-square statistics for statistical independence and contingency coefficients were selected taking into account the perceived importance in the injury severity analysis. Driver ejection and land use had both strong correlations and have been found significant in previous injury severity analyses.

Table 4-3: Driver Ejection and Injury Severity Cross Tabulation Table for Non-intersection Involvements

Ejected	Driver Injury Level					Total	Percent
Frequency (Row Percent)	PDO	Possible Injury	Evident Injury	Incapacitating Injury	Fatal Injury		
No	26305 (65.32)	7467 (18.54)	4505 (11.19)	1773 (4.40)	223 (0.55)	40273	96.41%
Total or partial	128 (8.53)	221 (14.72)	573 (38.17)	407 (27.12)	172 (11.46)	1501	3.59%
Total	26433	7688	5078	2180	395	41774	100.00%
Percent	63.28%	18.40%	12.16%	5.22%	0.95%		

From the cross tabulation shown in Table 4-3, it can be seen that driver partial or total ejection contributed to 18.7% of the incapacitating injuries and 43.5% of the fatal injury cases. Meanwhile only 3.59% of all crashes had an ejection event, which suggests that ejection significantly contributes to driver severe injury. Even though this event is technically a post-crash event, its inclusion in the model will measure proxies of crash precursors that lead to ejection, for example, the seat belt use. The nature of this relation for high-speed multilane roads can be an important topic for further research due to the significance of this variable in all the injury severity models.

Table 4-4: Urban Land Use and Injury Severity Cross Tabulation Table for Non-intersection Involvements

Urban_ID	Driver Injury Level					Total	Percent
Frequency (Row Percent)	PDO	Possible Injury	Evident Injury	Incapacitating Injury	Fatal Injury		
No	12163 (58.29)	4005 (19.19)	2904 (13.92)	1466 (7.03)	327 (1.57)	20865	49.95%
Yes	14270 (68.25)	3683 (17.61)	2174 (10.40)	714 (3.41)	68 (0.33)	20909	50.05%
Total	26433	7688	5078	2180	395	41774	100.00%
Percent	63.28%	18.40%	12.16%	5.22%	0.95%		

The urban land use was another significant association with the driver injury severity for multilane arterial roads. The cross tabulation in Table 4-4 shows that while the involvements were almost equally distributed between urban and rural areas, 82.8% of the fatal injury involvements and 67.2% of the incapacitating injuries occurred in rural areas. The land use classification is an important design parameter for arterial roads, which tend to have different design standards, such as drainage, shoulder and lane widths for rural and urban roads. Also, traffic speed is generally higher in rural sections than in urban sections. These important differences will likely have an impact on the crash mechanism and its outcome.

Table 4-5: Variables by Degree of Association to Injury Severity for Non-intersection Crashes

Degree of Association	Driver-related	Environment characteristics	Roadway characteristics
Strong	Ejected Speeding Alcohol Drugs Crash Damage (Vehicle) First Safety Equipment Vehicle Type Driver Action First Contributing Cause First Harmful Event Vehicle Movement Race	No strong relationships	Off/On Roadway Urban / Rural Traffic way Character Location Type Speed Limit Group Type of Shoulder Number of Lanes Location on Roadway
Weak	Gender Driver Age Group Vehicle Fault Code	Day of Week Time Group Road Surface Condition Bad Weather Lighting Condition	Undivided Highway

The stronger associations include only driver-related and roadway characteristic variables (see Table 4-5). In contrast, the weaker associations are mostly related to the environmental characteristics at the time of the crash. This does not negate the influence of the environment on crash occurrence, but it is understood that these variables may interact with the others to affect the injury severity. For each one of these strata, general association CMH tests were used to determine the best candidates for interaction among the weaker associations shown in Table 4-5. Driver's gender and vehicle fault code (driver at fault) were included in this analysis and are suspected to be confounders.

4.1.1.2 Non-intersection Crashes in Multilane Non-limited Access Roads

The group of non-intersection crashes (41,774 driver involvements) was further stratified by the road types. Urban and rural roads were classified by the number of lanes (two-three, four or more), speed limit (less than 40 mph and greater than 40 mph) and geometry of section (straight or curve). The subgroups are then examined for differences in their relationships with injury level. The group of crash involvements used in this section is summarized in Table 4-6. The analysis for two-three lane roads is not included here since these are not a main part of the investigation. The analysis was applied for roads with four or more lanes in rural/urban area and straight/curved sections. The resulting groups show the low driver crash involvement frequency for the curves of arterial roads. A total of 31,686 non-intersection driver crash involvements during 2004 were obtained for the four or more lane roads (no interstate or expressway).

Table 4-6: Non-intersection Driver Crash Involvements by Road Characteristics

	4 or more Lanes				**Total**	**Percent**
	Rural		**Urban**			
Road section horizontal alignment	**<40 mph**	**>=40 mph**	**<40 mph**	**>=40 mph**		
Straight	817	13374	4323	12062	**30576**	**96.50%**
Curve	30	599	156	325	**1110**	**3.50%**
Total	**847**	**13973**	**4479**	**12387**	**31686**	**100.00%**
Percent	**2.67%**	**44.10%**	**14.14%**	**39.09%**		

The general association (CMH) analysis by road characteristics was performed. If the p-value of the CMH statistic was less than 0.05, the variable has a statistically significant general association with driver injury severity when controlling for a given road characteristic. This non-parametric statistic is appropriate for the variables under study. The results are shown in Table 4-7, page 90, with marks on the variables with statistically significant general association. For roads segments with four or more lanes, speed limit below 40 mph, not on interstates or expressways, the rural roads show a lower frequency, as rural multilane roads tend to have higher speed limits than their urban counterparts. Only time group shows significant association with the other contributing factors when controlling for the weaker variables. This follows the trend of crashes in rural roads being related with more interactions from environment-related variables. For the urban low speed roads with four or more lanes, however, all the significant interactions are driver-related variables. In the lower speed roads, undivided highway has high p-values, which suggests no significant interaction by this road characteristic. Gender has an important interaction with the other factors.

Table 4-7: Variables with Significant Association between Injury Severity and Characteristics in Roads with 4 or More Lanes (Non-intersections)

Variable tested (general association) against injury severity	4 or more Lanes Straight Sections				4 or more Lanes Curve Sections			
	Rural		Urban		Rural		Urban	
	<40 mph	>=40 mph	<40 mph	>=40 mph	<40 mph	>=40 mph	<40 mph	>=40 mph
Time Group	X	X		X	Small sample	X	X	No general assoc.
Gender		X	X	X				
Undivided Hwy		X						
Weather		X		X				
Vehicle Fault Code		X	X	X		X		
Lighting Condition		X		X		X		
Driver Age Group		X	X					
Month								
Day of Week		X						
Road Surface Condition		X						

In the case of higher speed roads with four ore more lanes, rural areas display most of the possible interaction with significant p-values (<0.05). Even though the environmental variables are predominant, driver factors and the undivided highway variable are also significant. Higher speed rural (which are turning suburban with increased development) multilane roads tend to be important arterials which may carry local and non-commuter traffic. The mix of traffic and commercial development in urban areas increases the frequency and complexity of driver interactions, which may lead to a larger array of contributing causes that appear to interact for crashes occurring at these higher speed roads. The higher speed urban roads also display both environmental and driver-related variables with significant interactions; however, the urban straight sections seem to have less significant variables associated with injury severity when controlling for road characteristics. Some of the differences can be explained by the road and traffic characteristics. Urban roads are more likely to have median dividers and better drainage, reducing the interaction of undivided and surface condition variables. The driver age and day of week variables that were significant in the rural area and not in the urban area have more to do with traffic patterns and possibly car ownership, as explained previously.

Additional analysis results not shown in Table 4-7, page 90, provide additional insights into the relationships for certain variables. In the urban area, other environmental variables such as speeding and time group lost interaction significance when compared to the rural roads. It seems that speeding at night and crash severity is less notable in the urban area. Possible reasons include higher intersection density (shorter straight road sections), higher urban traffic and possibly enforcement at night when compared to rural areas. An additional detail found in the analysis is that alcohol/drug involvement in urban area roads was not significantly associated with injury severity when controlling for time group and lighting conditions. This may indicate a more uniform injury severity risk due to use of alcohol/drugs for drivers involved in crashes in the urban arterials. It remains to be seen how this relationship is related to driver behavior in urban areas. The relationships between the road characteristics and injury severity have important differences in rural and urban areas. Some of these variables, especially the alcohol/drug involvement proved to be challenging when developing the injury severity models.

For the rural higher speed arterials curved sections, time group, lighting condition, and vehicle fault code show interactions. There seems to be more crash occurrence in the rural multilane arterials, with more factors that seems to be interacting with the more strongly associated factors to injury severity. These results suggests that rural road conditions are different, perhaps more variant from section to section, and tend to have a greater influence over some important factor's association to driver injury severity. In additional three-way contingency tables it was found that safety equipment when controlling for gender seems to have a significant association with injury severity in many of the strata analyzed. This may indicate that driver behavior (choices) are generally affected by gender, although not as strongly as by age.

Table 4-8: Driver Involvements by Severity for Non-intersection Crashes on Low-speed Multilane Roads

	4 or more lanes low-speed (<40 mph) roads					
	Rural		Urban		Totals	Percent
Driver Injury Severity	Straight	Curve	Straight	Curve		
1 No Injury	558	19	3159	101	**3837**	**72.04%**
2 Possible Injury	169	7	687	26	**889**	**16.69%**
3 Non-Incapacitating Evident Injury	71	3	364	16	**454**	**8.52%**
4 Incapacitating Injury	17	1	104	13	**135**	**2.53%**
5 Fatal Injury	2	0	9	0	**11**	**0.21%**
Totals	**817**	**30**	**4323**	**156**	**5326**	**100.00%**
Percent	**15.34%**	**0.56%**	**81.17%**	**2.93%**		

To finalize the discussion of the preliminary analysis of non-intersection crashes, a brief look at the driver involvements by injury severity is warranted. For the curve sections of roads with four or more lanes, Table 4-8, page 92, indicates that drivers are less likely to be fatally injured in the lower speed roads. This is expected because of the lower operating and crash impact speeds. Under these conditions, severe crashes on curves are less likely to happen, which may partially explain the fewer variables with general association significance. In the lower speed urban roads, time group remained the only important interaction, while in the rural areas the sample size was too small to effectively analyze the general associations. Time group in the urban area may reflect the changes in traffic during the day or risky behavior during the night.

Table 4-9: Driver Involvements by Severity for Non-intersection Crashes on High-speed Multilane Roads

	high-speed (>=40 mph) multilane (4 or more lanes) roads					
	Rural		Urban		Totals	Percent
Driver Injury Severity	Straight	Curve	Straight	Curve		
1 No Injury	8252	269	8042	194	**16757**	**63.57%**
2 Possible Injury	2576	92	2238	54	**4960**	**18.82%**
3 Non-Incapacitating Evident Injury	1684	141	1299	48	**3172**	**12.03%**
4 Incapacitating Injury	722	70	442	24	**1258**	**4.77%**
5 Fatal Injury	140	27	41	5	**213**	**0.81%**
Totals	**13374**	**599**	**12062**	**325**	**26360**	**100.00%**
Percent	**50.74%**	**2.27%**	**45.76%**	**1.23%**		

In contrast with the lower speed multilane arterials, Table 4-9 shows a disparity between the crash involvements at curves on high-speed roads in urban vs. rural areas, where there is a notable difference in the proportion of severe crashes. The rural area crashes are not only more frequent, but tend to be more severe. This was also the case for two-three lane roads (analysis not shown here) where the number of fatal injuries on curves for high-speed roads was actually higher than for the 4 or more lane roads, even though the total number of crashes is particularly larger for the multilane arterials. This should be examined further in the future.

While the total driver involvements in urban and rural high-speed multilane roads are almost balanced, the severity of these involvements is not. The proportion of incapacitating injury in rural areas is much higher (63%) than in urban areas (37%). When considering curve sections the ratio of these involvements in rural vs. urban roads is almost 3 to 1. A similar situation occurs with the fatal injuries, confirming this trend. In contrast, the non-severe (levels 1-3) crashes are almost balanced for the rural and urban straight sections. The proportion of severe and fatal injury proportion (5.58%) was found to be higher than the 4.8% of a previous study using data from three counties in Central Florida (Abdel-Aty, 2003). It is closer to the statewide 2004-2006 average for total driver involvements (5.68%).

It must also be noted that when general CMH associations between driver action and other variables indicates that at least for one level, the driver action has a relationship with the injury severity of the driver. With additional contingency tables analysis, the driver action, phantom or hit and run, may be correlated to the crash severity. However, the data collection for these crashes may be affected by the difficulties associated with the investigation of a crash involving a hit and run or phantom driver. Data collection and accuracy issues will be presented in a later section.

4.1.2 Analysis of Intersection and Intersection-related Crashes

4.1.2.1 Intersection and Intersection-related Crashes in All Non-limited Access Roads

In this stage of the analysis, crashes occurring at intersections or within 250 ft of an intersection (influence area defined by DHSMV) for roads not classified as interstates or expressways are considered. The group of drivers involved in intersection crashes was higher than the non-intersection crashes occurring on state non-limited access roads during the year

2004. A total of 64,972 crash involvements were filtered from the 2004 crash records. Statistical tests of independence and measures of association for the possible contributing factors against the driver injury severity were computed using PROC FREQ, as previously mentioned.

Table 4-10: Preliminary Analysis of Intersection and Intersection-related Crash Driver Involvements (N=64,972)

Variable	Contingency Coefficient	DOF	Min (r,c)	Cramer's V	chi-square	p-value
Ejected	0.2517	4	2	0.2601	4395.2909	<.0001
Gender	0.1331	4	2	0.1343	1172.3599	<.0001
Vehicle Fault Code	0.1084	4	2	0.1090	772.5669	<.0001
Site_Location_ID	0.1082	4	2	0.1088	769.3757	<.0001
Rural/Urban	0.1035	4	2	0.1040	703.3809	<.0001
Alcohol-Drugs	0.0787	4	2	0.0790	405.3195	<.0001
Speeding	0.0666	4	2	0.0668	289.5686	<.0001
FL Resident	0.03	4	2	0.0300	58.4906	<.0001
Lighting Condition	0.0229	4	2	0.0229	34.1499	<.0001
Crash Damage	0.276	8	3	0.2030	5355.8425	<.0001
Vehicle Type (64,251)	0.2656	8	3	0.1948	4878.1981	<.0001
First Safety Equipment	0.2148	8	3	0.1555	3143.4986	<.0001
Driver Action	0.1213	8	3	0.0864	970.3235	<.0001
Location Type	0.0948	8	3	0.0673	589.1209	<.0001
Speed Group	0.0849	8	3	0.0603	471.7578	<.0001
Type of Shoulder	0.0634	8	3	0.0449	262.1445	<.0001
Road Surface Condition	0.0327	8	3	0.0231	69.5805	<.0001
First Contributing Cause	0.1369	12	4	0.0798	1240.2653	<.0001
Race	0.0548	12	4	0.0317	195.9059	<.0001
Time Group	0.0372	12	4	0.0215	90.2161	<.0001
Lane Group	0.0368	12	4	0.0213	88.2144	<.0001
First Harmful Event	0.172	16	5	0.0873	1979.6119	<.0001
Vehicle Movement	0.0993	32	5	0.0499	647.0668	<.0001
First Traffic Control	0.0798	16	5	0.0400	416.4241	<.0001
Driver Age Group	0.0751	28	5	0.0377	368.8244	<.0001
Number of Lanes	0.055	20	5	0.0276	197.3873	<.0001

The results of this analysis (see Table 4-10) illustrate some of the important differences when considering intersection crashes as`opposed to non-intersection crashes. The weather nor the off roadway variables are no longer significant at the 5% significance level. There were additional differences in the measures of association for the rest of the variables. There is a much stronger association of driver injury severity with gender and vehicle fault code than on the non-intersection crashes. Most roadway-related variables (off roadway, traffic way character, location type, speed limit group, type of shoulder, lane groups, location on roadway) loose association strength. . Meanwhile, two of the driver-related variables (speeding and alcohol-drugs) lost association strength. This suggests that the severity of driver injury resulting from crashes occurring at intersections tends to be more influenced by the drivers' actions than non-intersection crashes.

Table 4-11: Ejected and Injury Severity Cross Tabulation Table for Non-intersection Involvements

Ejected	Driver Injury Level					Total	Percent
Frequency (Row Percent)	PDO	Possible Injury	Evident Injury	Incapacitating Injury	Fatal Injury		
No	38849 (61.47)	12668 (20.05)	8502 (13.45)	2997 (4.74)	179 (0.28)	63195	97.26%
Yes or partial	191 (10.75)	321 (18.06)	685 (38.55)	470 (26.45)	110 (6.19)	1777	2.74%
Total	39040	12989	9187	3467	289	64972	100.00%
Percent	60.09%	19.99%	14.14%	5.34%	0.44%		

The statistics from Table 4-11 indicate that ejection events at intersection or intersection-related crashes (2.74%) are proportionally lower than for the non-intersection crashes (3.59%). When the ejection event occurs, intersection or intersection-related crashes injury outcomes comprise 38% and 13.5% of the fatal and incapacitating injuries, respectively. This suggests that in general crashes occurring at or near intersection are les likely to produce driver ejection and less likely to result in severe injury to those who are ejected from the vehicle. This comparison points out that crashes on segments are more likely to result in ejection, perhaps due to the single vehicle off-roadway crashes, which tend to be severe. The differences in crash mechanisms at segments and intersections are evident; however, the ejection event is still a very important factor that is present in 18.3% of all of the involvements under analysis (intersection and non-intersection).

Table 4-12: Urban Land Use and Injury Severity Cross Tabulation Table for Non-intersection Involvements

Urban_ID	Driver Injury Level					Total	Percent
Frequency (Row Percent)	PDO	Possible Injury	Evident Injury	Incapacitating Injury	Fatal Injury		
Rural	14589 (54.72)	5749 (21.56)	4254 (15.96)	1871 (7.02)	196 (0.74)	26659	41.03%
Urban	24451 (63.82)	7240 (18.90)	4933 (12.88)	1596 (4.17)	93 (0.24)	38313	58.97%
Total	39040	12989	9187	3467	289	64972	100.00%
Percent	60.09%	19.99%	14.14%	5.34%	0.44%		

The land use might be considered the most encompassing factor that deals with traffic behavior and road design. As shown in Table 4-12, the urban involvements are higher than the rural involvements. The rural involvements have a higher proportion of severe injury. While 41% of the involvements occur in rural areas, these represent 54% and 67.8% of the incapacitating and fatal injuries. Thus, similar to the segments, the rural sections of road present a serious trend of more severe injuries due to crashes. If the lower intersection density in rural areas is considered, there is some evidence of significantly higher severe crash rates at those rural intersections when compared to their urban counterparts.

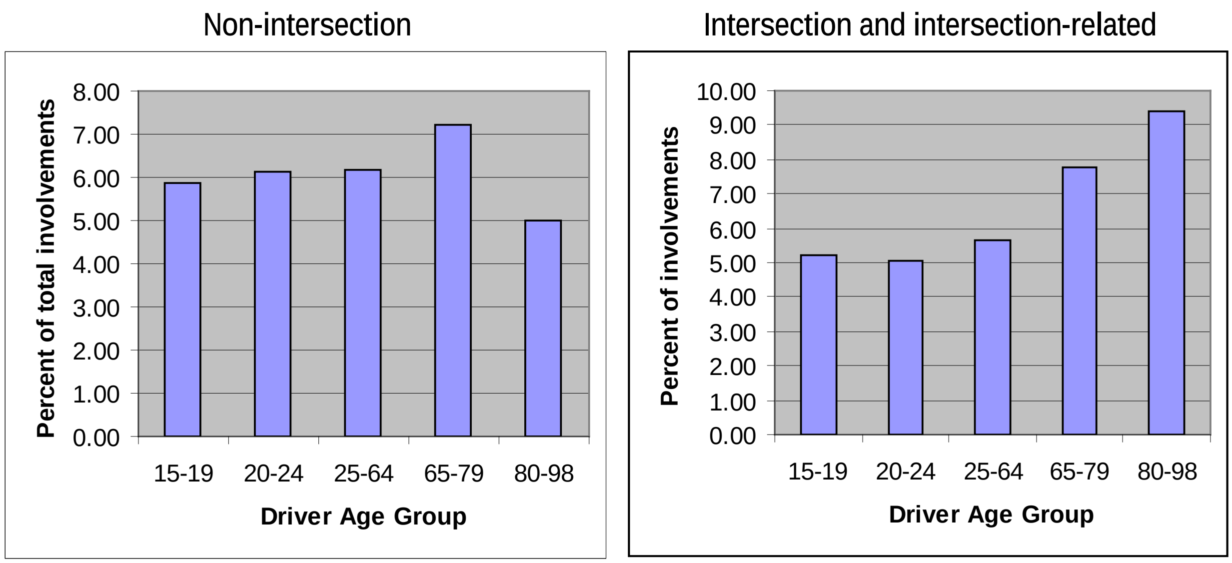


Figure 4-1: Severe Involvements in Non-limited Access Roads by Driver Age and Road Entity

In comparing the intersections and non-intersection environments, we consider two important factors of the drivers involved in crashes: age and gender. Figure 4-1 displays the distribution of severe involvements by driver age group divided in non-intersection and intersection events. While the non-intersection involvements exhibit an increasing proportion of severe injuries by driver age for the older drivers (65-79 years) and a reduction for the very old drivers (80 and above). On the other hand, for the intersection crashes the proportion of severe injuries for the younger drivers are lower than middle age drivers (25-64 years), but the older and very old drivers showed a significantly higher proportion. These statistics suggests that older drivers are the group a greatest risk of severe injuries. Meanwhile, the other age groups seem to be at greater severe injury risk in road segments.

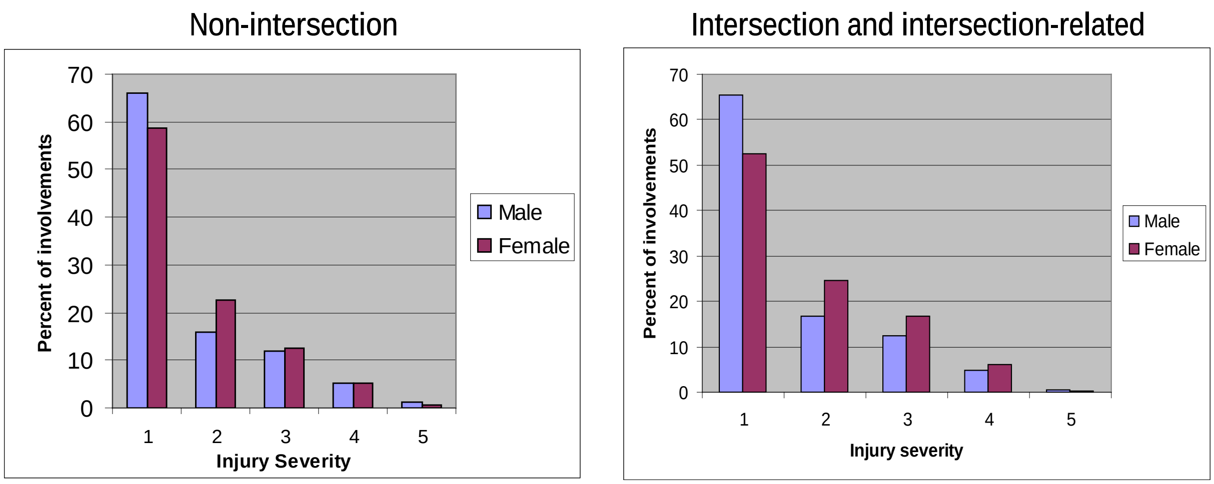


Figure 4-2: Severe Involvements in Non-limited Access Roads by Driver Gender and Road Entity

Driver gender affects the crash outcomes because of both the physiological and behavioral differences between males and females. The statistics shown Figure 4-2 indicate that female have more involvements in both types of crashes for the severe injury crashes. This preliminary analysis suggests the theory that females are more likely to suffer serious injury as a result of a crash event. Additional analysis in Chapter 5 will show these relationships applied exclusively to high-speed multilane roads.

The decreased association performance of the road-related factors should be analyzed in the context of the lack of intersection characteristics among the variables under analysis. As discussed in Section 2.4.1, additional intersection characteristics are needed to describe the safety performance of this road entity. Previous research by Abdel-Aty and Wang (2006) has shown that intersections present complex driving situations, especially in urban areas. Thus, drivers are required a greater degree of concentration and ability to traverse in a safely manner.

4.1.2.2 Intersection and Intersection-related Crashes in Multilane Non-limited Access Roads

As shown in Section 4.1.2.2, the crash involvements were stratified for additional CMH statistic analysis by the number of lanes, urban/rural area and speed limit (high / low). Intersection crashes were not classified as being on a curve/straight section, due to the inadequacy of this general description to a point location such as an intersection. Moreover, the traffic way character variable lost about 80% of its previous measure of association value, which confirms that it is no longer an important variable for this particular crash type. In addition, traffic volume has been proven to be an important factor to predict crash severity in intersections, as previously discussed. This variable was not available in the original crash records databases, but was derived from additional databases and will be incorporated into the analysis Section 4.2. The result of the general association analysis for multilane road character combinations is presented next.

The characteristics of intersections and its influence area in arterial corridors are more complex and require more information to describe the driver injury severity trends. In the case of four or more lanes road intersections in general more of the variables tested yielded significant interaction effect (association of two other variables when controlling for the variable being tested) This is probably due to more complex driving situations, as previously discussed. The most notable difference can be appreciated in Table 4-13, page 101, where the rural high-speed road intersection case shows a predominance of the driver-related variables interacting with the stronger variables (which are also mostly driver-related). These results suggest that driver behavior at these intersections may be of greater concern relative to other cases.

Table 4-13: Variables with Significant Association between Injury Severity and Characteristics in Roads with 4 or More Lanes (Intersection and Intersection-related crashes)

Variable tested (general association) against injury severity	4 or more Lanes Straight Sections			
	Rural		Urban	
	<40 mph	>=40 mph	<40 mph	>=40 mph
Alcohol Drugs	X	X	X	X
Speeding	X	X	X	X
Location Type	X	X		X
Type of Shoulder	X	X	X	X
Race		X	X	X
Driver Age Group	X	X	X	X
First Traffic Control	X	X	X	X
Lighting Condition			X	X
Time Group	X		X	X
Location on Roadway	X		*	X
Resident	X		*	X
Road Surface Condition	X			X

*** Indicates a marginally significant general association (0.05 ≤ p-value ≤ 0.20).**

In addition, some important differences between the intersection and non-intersection crashes for the high-speed roads can be seen. The format in Table 4-14, page 103, clearly shows the contrasts between the different road entity characteristics. For crash involvements in rural areas, the non-intersection straight sections had more statistically significant variables than for the intersections crashes, contrary to the urban areas. The most commonly significant variables were driver age group, lighting condition and time group. These variables were proved important in the injury severity models, including some interactions. The matrix also shows the importance of driver characteristics for intersections crashes; gender and driver fault are associated for non-intersection crashes.

With regards to roadway characteristics, location type (land use) and type of shoulder are associated with driver injury severity at intersections, while undivided highway is significant for rural non-intersection crashes. The type of shoulder could be an indirect effect of the intersection width (size), which will be discussed in Section 4.2. The higher operating speeds and design characteristics of rural segments certainly have an impact on the association between injury severity and highway divider. Road surface condition and weather are usually correlated and its use in the regression analysis was carefully tested to avoid multicollinearity. The day of week variable may correspond to trip purpose changes (weekday and weekend), and affect only straight sections in the rural areas.

Straight segments are more prone to higher operating speeds, increasing the risk of severe crashes. In addition, urban intersections have more complex relationships with injury severity, judging from the amount of variables with significant association. Another clue is that urban intersections are the only locations where a crash-related variable was found significant. This significance of the urban intersections was presumed when the research method was planned by including separate analysis by crash type, which is presented in Chapter 5. It is important to know what differences exist between intersection and non-intersection crashes that affect the driver injury severity.

Table 4-14: Variables with Significant Association between Injury Severity and Characteristics in Roads with 4 or More Lanes (Intersection and Intersection-related Crashes)

Parameter Group	Variable (general association) against injury severity	4 or more Lanes Straight Sections			
		Rural		Urban	
		Non-intersection	Intersection	Non-intersection	Intersection
Driver-Related	Driver Age Group	X	X		X
	Alcohol Drugs		X		X
	Speeding		X		X
	Race		X		X
	First Traffic Control		X		X
	Vehicle Fault Code	X		X	
	Gender	X		X	
	Resident				X
Roadway-Related	Lighting Condition	X		X	X
	Location Type		X		X
	Type of Shoulder		X		X
	Undivided Hwy	X			
Environment-Related	Time Group	X		X	X
	Road Surface Condition	X			X
	Weather	X		X	
	Day of Week	X			
Crash	Location on Roadway				X

Important differences between rural and urban crashes can also be inferred. Urban intersection environments have more possible factors associated with driver injury severity. In particular, the resident variable is exclusive of these intersections, suggesting that the degree of complexity or uniqueness in design is a possible reason for this significant association with the driver injury severity. On the other hand rural straight sections have a significant association between driver age and injury severity, but not the urban sections. Meanwhile lighting conditions did not show a significant association with injury severity in rural intersections. Other driver behavior, such as speeding and alcohol use seem to dominate the associations. Rural intersections did not show significant associations between environment-related variables and driver injury severity, but land use and type of shoulder were significant, perhaps differentiating between suburban and rural areas. However, rural segments have significant associations between the environment-related variables and driver injury severity. This points to the importance of weather conditions (related to visibility) in rural areas, similar to the freeway case.

Table 4-15: Multilane Road Intersection Crash Driver Involvements by Injury Severity (Intersection Crashes)

Driver Injury Severity	4 or more Lanes Roads					
	Rural		Urban		Totals	Percent
	<40 mph	>=40 mph	<40 mph	>=40 mph		
1 No Injury	1689	9805	7549	12447	**31490**	**59.70%**
2 Possible Injury	667	3855	2350	3752	**10624**	**20.14%**
3 Non-Incapacitating Evident Injury	473	2903	1473	2742	**7591**	**14.39%**
4 Incapacitating Injury	216	1207	457	938	**2818**	**5.34%**
5 Fatal Injury	25	113	34	48	**220**	**0.42%**
Totals	**3070**	**17883**	**11863**	**19927**	**52743**	**100.00%**
Percent	**5.82%**	**33.91%**	**22.49%**	**37.78%**		

The second important comparison between intersections and non-intersection crashes is the steady increase of the ratios of driver involvements between rural and urban high-speed intersections as the injury level increases. Table 4-15 shows the driver involvements by injury severity for intersection crashes. From the information presented, the rural to urban ratio for non-injury crash involvement is 0.78, but for incapacitating injury is 1.29 and fatal injury crash involvement it increases to 2.35. Another important observation for the high-speed crashes is that while the total number of crash involvements (37,810) is larger than those occurring at straight sections, the number of fatalities is lower for the intersection crashes, the intersection (and intersection-related) crashes resulted in almost double the number of incapacitating injuries. Special attention was given to these situations in the next stage of the investigation.

4.1.3 Distribution of Intersection and Intersection-related Severe Crashes

Preliminary analysis also investigated the frequency of severe crashes on a subset of crashes selected by their site location code. From the year 2004 data, after pre-screening for complete records, there were 12,487 intersections with at least one crash and 3,075 intersections with at least one severe crash. There are some important contrasts between the distributions, which are presented in Appendix A. It is important to point out that all of the distributions were found to be similar to the Poisson distribution, with high proportions of intersections with low crash counts. This observation needs to be confirmed by future research using the full population of intersections. This analysis was divided into the signalized and unsignalized intersections in the urban and rural areas. The number of intersections with at least one severe crash was higher for the signalized intersections, especially in urban areas. Among the unsignalized intersections, those in rural areas had more severe crash frequencies.

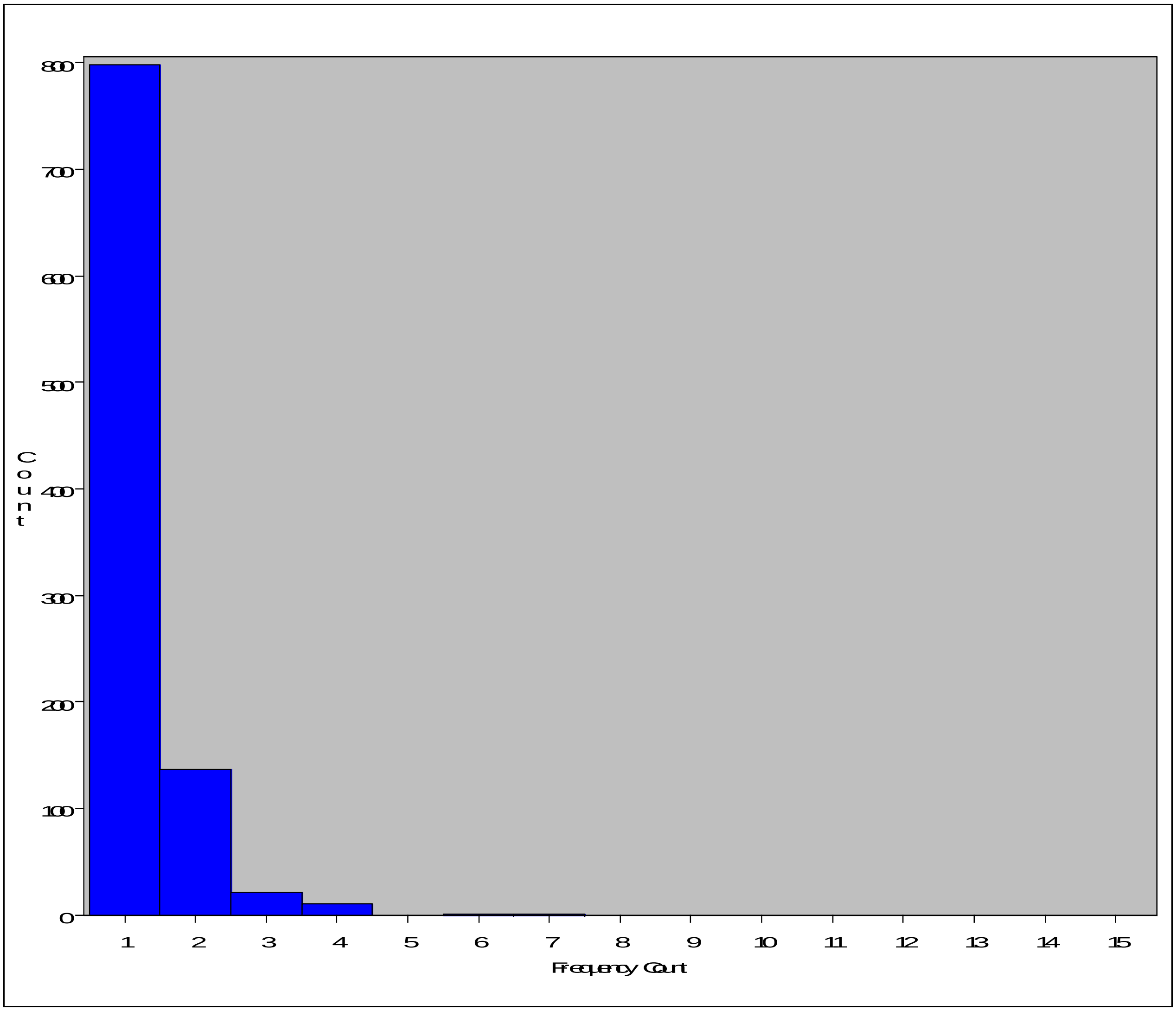


Figure 4-3: Severe Crash Counts for At (or Near) Urban Signalized Intersections on Multilane Arterials

Previous research in California has shown a tendency of higher total and injury crash frequencies for urban signalized intersections (Bauer and Harwood, 1996). Figure 4-3 shows the accident frequency distributions of driver crash involvements for urban signalized intersections in multilane arterials, which was higher than the other kinds of intersections. This graph suggests that close to 20% of the urban signalized intersections reporting severe crashes in 2004 had more than one crash. In addition, there are some smaller proportions of intersections with more than five severe crashes in 2004, which was not evident in the graph for the rural signalized intersections (see Appendix A).

This brings the question of exposure vs. number of locations. On one hand, there are more urban signalized intersections than rural; it is expected to have larger numbers of urban signalized intersections with at least one severe crash. On the other hand, the traffic volumes in rural intersections are generally lower and there are similar numbers of intersections with multiple severe crashes in the rural and urban areas. Additional research with intersection traffic volume data is needed in order to determine the levels of risk for severe crashes at the urban and rural signalized intersections on arterials.

The graph in Figure 4-4, page 108, clearly shows a smaller number of unsignalized intersections with one or more severe crashes than the urban or rural signalized intersections. These results seem to contradict the benefit of decreased injury severity in signalized intersections. Signalized intersection volume and crash warrants dictate that those with certain crash frequency thresholds would be candidates for a traffic signal. This selectivity affects the risk exposure of the signalized intersections. Regardless of the exposure, the number of intersection severe crashes represents a majority of the crashes at multilane arterials, as described in the next sections.

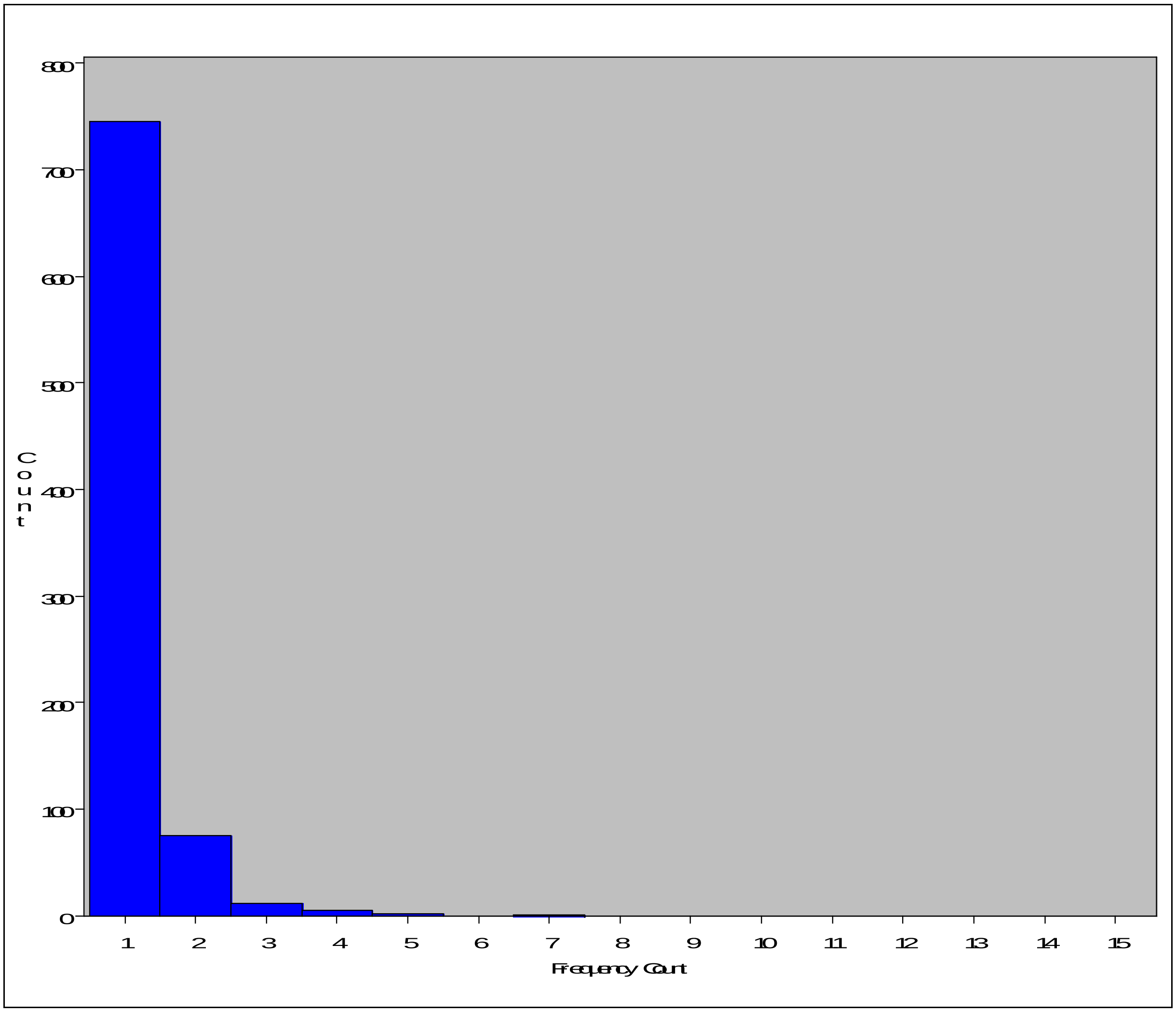


Figure 4-4: Severe Crash Counts for At (or Near) Rural Unsignalized Intersections on Multilane Arterials

This graphical representation analysis suggests a pattern of increased severe crash frequency on signalized intersections in urban areas compared to other kinds of intersections on multilane arterials. This analysis doe not pretend to determine which type of intersection is more risky, but rather to point out the differences between them in terms of severe crash occurrence. The differences in severe crash frequency patterns are important when making model building choices. In this one year analysis, the type of traffic control is an important factor in model building. It shall become apparent at the end of the analysis section.

4.2 Exploratory Regression Analysis of One-year Data by Driver Involvement

An exploratory regression analysis by road entity was performed using driver involvements for the 2004 crash data. Logistic regression models with a binary response (severe vs. non-severe driver injury) were developed in the exploratory stage. Six models resulted from the analysis: all involvements (OVERALL), all intersections (INTERS), signalized intersections (SIGNAL), unsignalized intersections (UNSIG), pure segment (PURE SEG), and pure segments plus unsignalized intersections (SEGMENT). An additional model was developed for single vehicles crash involvements at intersections (SIG 1VEH) but it had poor goodness of fit and was not further considered in this exploratory analysis.

For exploratory logistic regression analysis, additional variable data restrictions reduced the original sample size, as shown in Table 4-16, page 110. This table also shows the response profile for each subset of data, for which it can be seen that the lowest proportion of severe driver injury involvements are for the involvements occurring within *pure* segments. Pure segment involvements are those outside of the signalized or unsignalized intersection area of 250 ft or not related to its operation; following the site location definition, as explained previously. From this analysis, it is apparent that the proportions are not homogeneous, as demonstrated by the chi-square test p-value (<.0001). When considering all these models, the contingency coefficient (a measure of association) is 0.0870. Since these are not all independent samples, this measure only serves as a comparison tool to be used later on. The larger proportions of severe crashes and higher frequencies of severe injuries for the intersection models reflect that in multilane arterials the intersections play a preeminent role in the driver injuries.

Table 4-16: Response Profile for Severity Model Data

	OVERALL	INTERS	SIGNAL	SIG 1VEH	SEGMENT	PURE SEG.	UNSIG
Total involved	60221	32651	18956	19339	36447	24332	15102
Severe injuries	3550	2069	1046	1076	2312	1324	1093
% Severe injuries	5.89%	6.34%	5.52%	5.56%	6.34%	5.44%	7.24%
% Difference (vs. pure segment)	**8%**	**16%**	**1%**	**2%**	**17%**	**BASE**	**33%**

Test of homogeneity p<.0001, Contingency Coefficient=0.0870

The proportions shown in Table 4-16 suggest a trend toward higher crash injury severity for involvements occurring at unsignalized intersections, as well as lower proportions for the signalized intersections and pure segments. It was expected that the signalized intersections would affect the crash occurrence in a multilane arterial corridor due to the high concentration of these, especially in urban areas. The preliminary analysis has shown that the segment crash involvements injury severities are affected by different factors than those related to signalized intersections. However, some questions remain:

- How can the best models for severe crash involvements for drivers in a multilane arterial be obtained? Is there a need or enough information for including interaction terms not included in previous research?
- Do signalized intersections and pure segments in fact constitute location and traffic control combinations worthy of separate modeling?

- Is it better to model some of these combinations together because their factors influencing severity (and corresponding countermeasures) are similar?

The distribution of the driver injury severity of the high-speed multilane arterial crash involvements follows the trends already discussed in the preliminary analysis. A total of 60,221 crashes (from a complete dataset with a total of 106,746 crash records) were found in the prepared data for the year 2004 and used in the overall model. The driver injury severity distribution for the crash involvements utilized is presented in Figure 4-5. Having analyzed the overall crash involvements and following the objective of finding possible differences between crashes at different locations and traffic controls, the next step was to analyze intersection crashes.

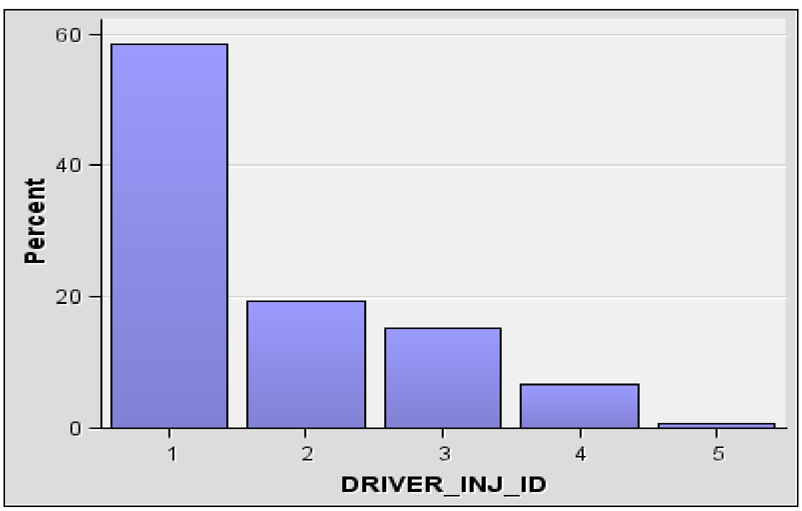


Figure 4-5: Driver Injury Severity Distribution for All Involvements on High-speed Multilane Arterials

The comparison of the driver injury distribution for the different road entities and combinations used in the six models are paramount to understand the relative importance of the results. The graph in Figure 4-6, page 112, shows that the injury distribution in the unsignalized intersections is higher for the severe and lower for the non-severe injuries. In Figure 4-7, page

112, the pure segments denote lower severe injury proportions. The segment model (combination of pure segments and unsignalized intersections) seemed homogenous with the overall sample in terms of the injury severity distributions.

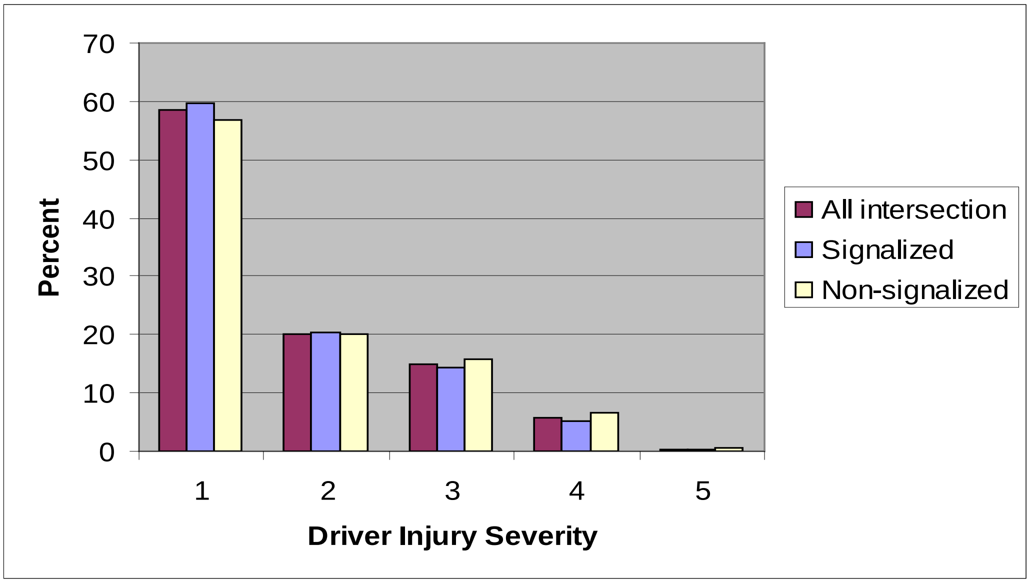


Figure 4-6: Driver Injury Severity for Involvements At (or Near) Unsignalized Intersections

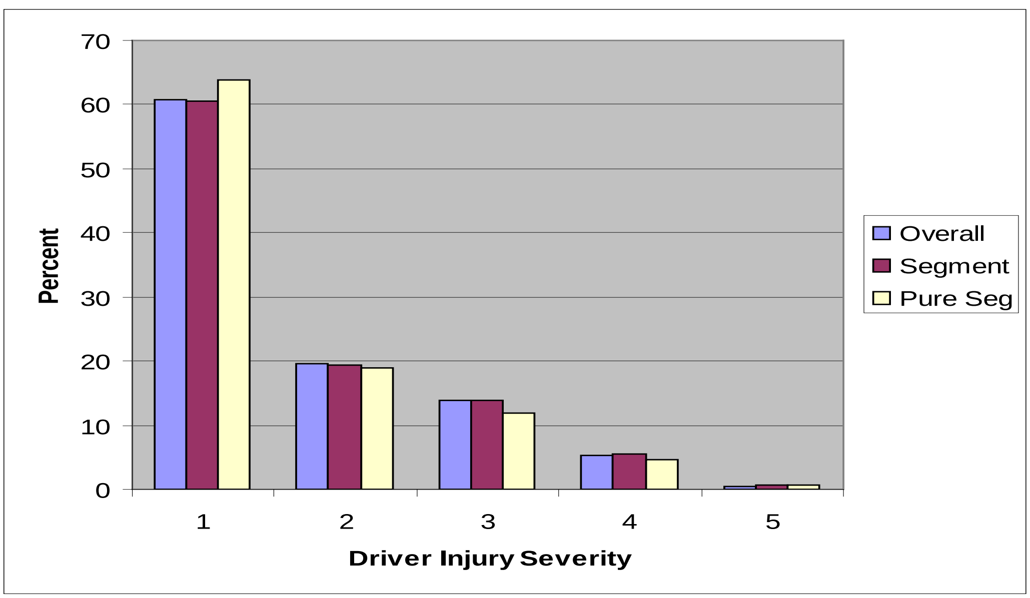


Figure 4-7: Driver Injury Severity for Involvements on Pure Segments

4.2.1 Injury Severity Model Building and Interpretation

A total of six models were calibrated using binary logit regression following the analysis tree presented in Section 3.6. An additional model was developed (but not used in the analysis due to the small sample size) for single vehicle crash involvements at signalized intersections after repeated lack of fit and numerical problems when single and multiple vehicle crash involvements at or related to intersections were combined into one severity model. This suggests that single and multiple vehicle crashes at the intersection area be considered as an important factor when using more advanced statistical methods. In previous research by Nasaar et al. (1994) using a nested logit model to predict driver crash injury severity, the differences of single and multi-vehicle crashes were apparent. In that investigation, three separated models were calibrated according to the crash situation: single-vehicle, two-vehicle and multiple vehicle crashes. The six models were developed and their goodness of fit was assessed as discussed in forthcoming Section 4.3. Three models were selected for their better statistical qualities: unsignalized intersections, signalized intersections (multiple vehicle movements) and pure segments. The results of these three models will be discussed next.

The response value was coded one (1) for severe (incapacitating or fatal) injury or zero (0) otherwise. The probability modeled was y=1, making the odds ratios easily interpreted as the probability of a severe injury. After model building, the significant coefficients and their values were examined for scientific validity. Their particular values were not as important as their signs (increasing or decreasing risk of severe driver injury) for this stage of the project. The values obtained can be later compared with the multiyear model data. Next, the models were compared in terms of their significant variables and their validity. Some positive results in terms of coefficients observed in the three selected models over the three remaining models were found during model building. Some of these findings were about variables that were not present in the selected models, but showed up in some of the other models with less favorable goodness of fit. The following summarizes these findings:

- Some confounding interactions, such as driver fault or aggressive driving with the collision type or traffic control, were not found significant in the selected models. However, the always important speed-related, driver age and gender, were significant factors in the selected models.
- Vehicle movements, which may cause confounding effects with the first harmful events, were not significant in the models selected, contrary to other competing models.
- Crash lane and district (region) variables were not found to be significant for the selected models. These variables did not contain a high interpretative value and had some confounding effects on the larger models.

4.2.1.1 Driver-related Significant Factors

Once these positive differences were recognized, the focus is on comparing the relationships between the significant factors for the three models. First, the driver-related effects were compared. Table 4-17, page 115, shows the significant factors present in each model and the coefficient interpretation (Increased or Decreased Severe Injury Risk). The increased risk is defined as odds ratio above 1.0, the decreased risk with an odds ratio less than 1.0. The blank cells indicate that the variable is not present in the model. Variables in light yellow indicate marginal individual significance levels (0.05<=p-value<=0.1). Yellow cells indicate that this variable (or level) was found non-significant on the basis of the coefficients (Maximum Likelihood Estimates) Wald chi-square analysis (p-value>0.1). Some of the variable levels labeled *other* are expected to be non-significant, as various different values are combined in these categories due to their sparseness.

Table 4-17: Driver-related Significant Factors in Injury Severity Models with Increased Odds Ratio (Positive Coefficient) or Decreased Odds Ratio (Negative Coefficient) Compared to Severe Injury Odds Ratio=1

Model / Parameter	PURE SEGMENT	SIGNALIZED INT.	UNSIGNALIZED
Driver Age 15-19 (vs. 25-64)	Decreased	Decreased	Decreased
Driver Age 20-24 (vs. 25-64)	Decreased	Decreased	Decreased
Driver Age 65-98 (79) (vs. 25-64)	Increased	Increased	Increased
Driver Age 80-98 (vs. 25-64)		Increased	
Female (vs. male)	Increased	Increased	Increased
Driver at Fault	Decreased	Decreased	Decreased
Aggressive Driving (vs. no improper action)	Increased	Increased	
Other Contributing Cause (vs. no improper action)	Increased	Decreased	
Aggressive Driving			Increased
Aggressive_*Rear_End	Decreased		Decreased
Aggressive_*Head_On	Decreased		
Alcohol/Drugs Involved	Increased		Increased
Speeding			
Speed Ratio 0.91-1.26 (vs. 0-0.9)	Increased		
Speed Ratio 1.26-2 (vs. 0 - 0.9)	Increased		
Speed Ratio 1.26-2 (vs. 0-1.26)			Increased
Estimated speed	Increased	Increased	Increased
Seat Belt (vs. none)	Decreased	Decreased	Decreased
Safety Helmet (vs. none)	Increased	Increased	
Other Safety Equipment (vs. none)	Increased	Increased	Increased
Ejected	Increased	Increased	Increased

The results shown in Table 4-17 indicate that all models have factor coefficients that agree with the scientifically expected effects on driver injury risk. Younger drivers are less likely to be severely injured than middle or older drivers. Older drivers, as expected, are the group most likely to sustain severe injury. Females are found to have increased risk of severe injury than males. Driver group characteristics proved to be important factors, although driver behavior factors have the most impact on crash injury severity, as shown in the previous section.

Drivers at fault are found to have less probability of severe injury than those not at fault. One additional variable was derived from the contributing cause field in the crash database. Aggressive driving was defined in the preliminary analysis as any of the following driver actions (as defined by FDHSMV): Careless Driving, Failed To Yield Right-Of-Way, Improper Lane Change, Improper Turn, Followed Too Closely, Disregarded Traffic Signal, Exceeded Safe Speed Limit, Disregarded Stop Sign, Improper Passing, Exceeded Stated Speed Limit, Disregarded Other Traffic Control. Aggressive driving (vs. no improper driver action) was found to be a factor that increased the severe injury risk for the pure segment and signalized intersection models. For the unsignalized intersection model a similar increased risk effect was observed in the aggressive driving binary (aggressive driving vs. no aggressive driving). This *dummy* variable was required to avoid confounding effects that affected other related variables and interactions. Meanwhile, the interactions of aggressive driving with rear-end and head-on collision types resulted in decreased risk of a severe driver injury. The interaction with the head-on crash type suggests that other unsafe driver behavior is at work to make them more likely to result in severe injuries. Kim et al. (1995) found that driver behavior such as alcohol or drug use and lack of seat belt use greatly increased the odds of more severe injuries for head-on crashes. In the case of rear-end crashes, it seems that aggressive driving does have an effect on rear-end crash occurrence, which generally result in less severe injury.

The speed ratio variable indicates the degree to which a driver's estimated speed is below at or above the speed limit. This estimated speed to posted speed ratio was classified into two (0-

1.26, 1.26-2) or three categories (0-0.9, 0.91-1.26, and 1.26-2). In the pure segment and unsignalized intersection models, the significant coefficients suggested an increased severe injury risk for those drivers over the speed limit by at least 26%. However, the speed ratio factor was not significant for the signalized intersection model. In the pure segments, the changing lanes maneuvers and the left turns from minor roads (gap acceptance in unsignalized intersections) may be major causes that explain the significance of this factor and that are not present with the traffic signal. Seat belt usage has been repeatedly proven to be a major factor reducing injury severity, and it is present in all three models. Finally, being ejected (no use of restrain devices) has an opposite and magnified effect: a much larger risk of being severely injured than the reduction shown in the seat belt factor. There is a correlation between the two, for which it may be said that there is a great overall positive effect of using the seat belt and other restrain devices.

4.2.1.2 Crash-related Significant Factors

The crash-related effects are shown in Table 4-18, page 118. The major effects of increasing driver injury severity (compared to rear-end crashes) are the crash types: head-on, angle, fixed-object, and left turn, as expected. The overturned crashes may be correlated to the fixed-object-related crashes; this is supported by the non-significance of this variable in the signalized intersection driver crash severity model. Pedestrian and bicyclist crashes are also expected to contribute to decreased risk of driver severe injury, which is our response variable. Vehicle type was found to be significant for the pure segment and unsignalized intersections models. The total number of vehicles involved is also significant (increasing risk), but only in the pure segment model (but not in the competing segment plus unsignalized intersections model).

This is expected, as multiple (three or more) vehicle crashes (such as cross median) are severe and more likely to occur on open road segments. The point of impact variable relates the location of initial impact on the vehicle and driver injury severity. As expected, frontal crashes and those by the driver's side are more likely to increase the driver's injury severity.

Table 4-18: Vehicle and Collision-related Significant Factors in Injury Severity Models with Increased Severe Injury Odds (Positive Coefficient) or Decreased Severe Injury Odds (Negative Coefficient)

Model / Parameter	PURE SEGMENT	SIGNALIZED INT.	UNSIGNALIZED
Head-On collision (vs. Rear-end)	Increased	Increased	Increased
Angle (vs. Rear-end)	Increased	Increased	Increased
Left Turn (vs. Rear-end)	Increased	Increased	Increased
Right Turn (vs. Rear-end)		Decreased	
Sideswipe (vs. Rear-end)	Decreased	Decreased	Decreased
Pedestrian (vs. Rear-end)	Decreased		Decreased
Bicyclist (vs. Rear-end)	Decreased		Decreased
Fixed Object (vs. Rear-end)	Increased	Increased	Increased
Overturned (vs. Rear-end)	Increased		Increased
Other Collision (vs. Rear-end)	Increased	Increased	Increased
Pedestrian and Bicyclist (vs. Rear-end)		Decreased	
Bus/Trucks (vs. Passenger car/ van)	Decreased		Decreased
Bike/motorcycle (vs. Passenger car/ van)	Increased		Increased
Total Number of Vehicles	Increased		
Point of Impact (Front Right vs. Front)	Decreased	Increased	Increased
Point of Impact (Back Right vs. Front)	Decreased	Decreased	Increased
Point of Impact (Back vs. Front)	Increased		Decreased
Point of Impact (Back Left vs. Front)	Decreased		
Point of Impact (Front Left vs. Front)	Increased	Increased	
Point of Impact (Other vs. Front)	Increased	Decreased	Decreased
Point of Impact (Back + Back Left vs. Front)		Increased	Decreased
Off Roadway crash	Decreased		Increased
Off-Road*Speed Ratio 0.91-1.26 (vs. 0-0.9)	Decreased		
Off-Road*Speed Ratio 1.26-2 (vs. 0 - 0.9)	Decreased		

In addition, there are differences in most coefficients between the unsignalized and pure segment models, which may indicate underlying important differences in crash mechanisms that should be further investigated. First, the front right point of impact is only significant for the signalized intersection model. Second, the off-roadway crash variable has a well-defined increasing severity risk coefficient for the unsignalized intersection model, but conflicting coefficients in the pure segment model. Also, the pure segment coefficients are marginally significant or non-significant for the interaction between the off-roadway crashes and the speed ratio. This may be the result of selectivity of the 250 ft radius as defined for intersection-related crashes in Florida used in the crash database, regardless of its traffic control type, physical size or traffic volume. However, every other indicator is supporting the validity of these models and no evidence of misspecification was found.

4.2.1.3 Roadway-related Significant Factors

The road-related effects (see Table 4-19, page 120), although less in quantity than the driver-, vehicle- and crash-related factors, are important in this severity model. Recall that one of the objectives of this analysis was to find some of the important road characteristics factors not traditionally reported in severity analysis. In addition, there might be correlations between some of the road characteristics and driver-, crash-, vehicle- or environment-related factors. These correlations may prove to be important in this study. The traffic control variable, although limited, suggests that stop control is better than no control at all. Crash involvements on high-speed multilane arterial roadways with urban land use consistently have lower driver injury severity risk than their rural counterparts, as expected.

Table 4-19: Roadway-related Significant Factors in Injury Severity Models with Increased Severe Injury Odds (Positive Coefficient) or Decreased Severe Injury Odds (Negative Coefficient)

Model / Parameter	PURE SEGMENT	SIGNALIZED INT.	UNSIGNALIZED
Stop Sign Control (vs. None)			Decreased
Other Traffic Control (vs. None)			Increased
Urban (vs. Rural)	Decreased	Decreased	Decreased
SPEED_LIMIT_ID3 (55-70 mph vs. 40-50 mph)		Increased	
Speed Limit 50 mph (vs. 40-45 mph)	Increased		Decreased
Speed Limit 55 mph (vs. 40-45 mph)	Decreased		Increased
Speed Limit 60-70 mph (vs. 40-45 mph)	Increased		Increased
***adt* per Lane (in thousands)**	Decreased		Decreased
Shoulder Width (<3.5 vs. =>3.5)	Increased	Increased	Increased
Average Truck Factor (%)		Increased	
Surface Width (ft)		Increased	

However, the speed limit effect is not consistently represented. The unsignalized and signalized intersections models show increased risk with increased speed limit (55 mph or higher vs. base 40-45 mph). The different variable categories reflect the uniqueness of each model, as optimized using data mining techniques. The pure segment model, however, is not consistent with the coefficient values nor are they completely significant. Selectivity (influence area) might be an issue, but it is more likely that the interaction between off-roadway crash involvements and the speed ratio might suggest that driver behavior is a major factor in the most severe crashes on pure segments, rather than the higher speed road design features.

For the intersection models, increased intersection space size on the major road (larger surface and shoulder width) contributes to higher severe driver crash injury involvements. This inference considers that fact that minor road crash involvements within 250 ft of all intersections are included in this analysis. Therefore, the longer distance to cross the intersection increases the risk (vehicle exposure) to both angle and left crashes of the minor road vehicle with a vehicle traveling on the high-speed multilane arterial. The relation of intersection size with high crash frequency and severity with at least one major type of intersection (signalized four legs, two-way) has been already shown in previous research by Abdel-Aty et al. (2006).

In the case of the pure segments, shoulder width has been shown to be a positive factor in speed of the traffic flow. However, neither speed limits nor speed ratio interactions with shoulder width were found to be significant in this model. Other underlying relationships that may become apparent in the spatial analysis are the rural vs. urban area road characteristics, such as intersection (and driveway) density and the presence of right turn auxiliary lanes; which may further describe this relationship.

Finally, traffic factors are also significant in the three models. Increased traffic volume per lane, which is a normalized measure of exposure, was found to decrease the chance of severe driver injury severity given a crash involvement. Previous research has shown that, under higher traffic volume conditions, the frequency of crashes tends to increase but also the average speed of traffic decreases, thus resulting in lower injury severity. Also, an increased average truck volume percent contributes to more severe driver crash involvements at signalized intersections.

4.2.1.3 Environment-related Significant Factors

The environment-related factors are shown in Table 4-20, page 122. It appears that the pure segment driver injury severity is most affected by the environmental variables. However, some of these variables levels are not significant. Time of day is significant for the pure segment model and shows increased risk for drivers during nighttime, which follows previous empirical relationships. The most consistent is the lighting condition variable, which is significant for the three models and shows the negative effect (increased driver injury severity) of lack of street lighting at night. There is also a positive effect for the street lighting, but it is not significant for the pure segments models, perhaps due to a smaller sample of lighted road segments when compared to intersections.

Table 4-20: Environment-related Significant Factors in Injury Severity Models with Increased Severe Injury Odds (Positive Coefficient) or Decreased Severe Injury Odds (Negative Coefficient)

Model / Parameter	PURE SEGMENT	SIGNALIZED INT.	UNSIGNALIZED
Dark- with street lighting (vs. daylight)	Decreased	Decreased	Decreased
Dark- no street lighting (vs. daylight)	Increased	Increased	Increased
Head_On*Dark- with street lighting (vs. daylight)	Increased		
Head_On*Dark- no street lighting (vs. daylight)	Increased		
Sideswipe*Dark- with street lighting (vs. daylight)	Increased		
Sideswipe*Dark- no street lighting (vs. daylight)	Decreased		
Time Group (12AM-6AM vs. 6AM-Noon)	Increased		
Time Group (6PM-Midnight vs. 6AM-Noon)	Increased		
Time Group (12PM-6PM vs. 6AM-Noon)	Decreased		

4.2.2 Variable Relative Significance Analysis

The variable relative significance analysis results (see Table 4-21, page 123) show that the Pure Segment crash involvement driver injury severity are most influenced by the driver-related variables (gender, driver fault, contributing cause, safety equipment, ejected) and the first harmful event (collision type), followed by the average traffic volume per lane and point of impact variables.

Table 4-21: Relative Variable Significance in Exploratory Injury Severity Models

Contributing Factor	Major (p<.001)	Moderate (.001<=p<.01)	Minor (.01<=p<.05)
Driver Age Group	All		
Gender	All		
Driver Fault	All		
Aggressive_Driving	All		
Estimated_speed	All		
Safety_Equipment	All		
Ejected_ID	All		
First_Harmful_Event	All		
Point_of_Impact	All		
Urban_ID	Unsign, Seg	Sign	
SPEED_LIMIT	Unsign, Sign		Seg
Shouder Width	Sign	Unsign	Seg
Lighting_Condition		Unsign	Seg, Sign
Rear-end*Aggressive	Seg	Unsign	
Alcohol_Drugs_ID	Unsign	Seg	
Speed_Ratio	Seg	Unsign	
VEH_TYPE_GROUP	Seg, Unsign		
Off_Roadway_ID		Unsign	Seg
ADT_Per_Lane	Seg	Unsign	
Aggressive*Head_On			Seg
Number_of_Vehicles		Seg	
Ped*Num_Vehicles		Seg	
Off_Roadw*Speed_Ratio			Seg
Head_On*Lighting_Cond.			Seg
Sideswipe*Lighting_Cond.			Seg
Time_Group	Seg		
First_Traffic_Contro	Unsign		
AVG Truck Factor			Sign
SURFACE WIDTH			Sign
Road_Surface_Condition			Sign

All= all models, Unsign= Unsignalized, Sign=Signalized, Seg=Pure Segment.

In contrast, the unsignalized intersection model is more influenced by safety equipment, followed by collision type and some of the driver-related variables (ejected, age group, gender), but in this case age being more important than gender, which agrees with previous research that point to the differences in ability of drivers by age groups required to accept gaps (minor road drivers) and major road drivers reaction to vehicles crossing from the minor road at unsignalized intersections. The speed factor relative significance points to the issue of the fixed-object crash significance mentioned earlier as well as to the gap acceptance for the vehicle in the minor road. Other elements like point of impact, vehicle type; alcohol/drug use, aggressive driving, and speed limit have a moderate effect, which suggests that severe crashes at unsignalized intersections are more influenced by road features when compared to the pure segment involvements.

Finally, the signalized intersection crash involvements are more influenced by the collision type, safety equipment, and ejected variables, as the other models. However, the driver's estimated speed is the next most influential factor, followed by the gender and age, which are more influential as a pair compared to the other models. The gender and age influence points to a more balanced correlation, which may suggest that there are more complex issues at work in the signalized compared to the unsignalized intersections.

In terms of environment-related variables, the Pure Segment model is more influenced by the Time Group variable, which may be correlated to driver behavior. Nevertheless, it pinpoints to some of the most effective methods to reduce severe crashes at the arterial segments. Lighting condition has a minor relative significance; only in the unsignalized intersection models it has moderate relative significance. The street lighting factor is an important design feature that is less frequently found on unsignalized (smaller) intersections than at signalized intersections.

In conclusion the most important variables for the crash involvements driver injury severity are: collision type (harmful event), safety equipment and ejection, gender, age, estimated speed (signalized intersections), driver fault, contributing cause (and aggressive driving), and point of impact. Other major and moderately influential variables include speed ratio (unsignalized), alcohol/drugs and speed limit (signalized and unsignalized), and traffic volume (unsignalized). Lighting condition is more influential in the unsignalized intersection model, while shoulder width is more so in the signalized model.

4.3 Exploratory Models Reliability Comparison

The six injury severity models were developed and their goodness of fit was compared in order to select the best models that explain all the crashes occurring on high-speed multilane arterial corridors. Models with larger groups of crashes were compared against smaller, more disaggregated models. The results summary of the model assessment (see Table 4-22, page 126) shows that a model based on all of the crash involvements at intersections, although has an acceptable classification accuracy, fails the Hosmer-Lemeshow decile of risk goodness of fit test ($p\text{-value}<0.05$). The convention of this test is that the larger the p-value, the better is the model fit. Another measure for comparison is the difference of the deviance (equal to the SSE when computed for a linear regression), which measures the improvement in the classification accuracy of the response variable. The AIC value is use due to its asymptotic efficiency.

Table 4-22: Competing Models Assessment Summary

GOF Parameter	OVERALL	INTERS	SIGNAL	SEGMENT	PURE SEG	UNSIG
Number of Variables	29	26	16	23	27	20
Degrees of freedom	60	53	34	49	54	40
Marginally significant levels	3	3	2	5	1	2
Non-significant levels	8	6	4	7	14	7
AIC	22802.63	13382.04	7301.98	14145.56	8201.96	6707.86
Hosmer-Lemeshow p-value	0.4419	0.0475	0.2594	0.9837	0.5474	0.5354
c value(area under ROC curve)	0.782	0.765	0.738	0.799	0.82	0.773
Percent Concordant	77.6	76	73.1	79.5	81.4	76.8
Deviance	22680.63	13274.04	7231.98	14045.56	8091.96	6625.86

Graphical methods were used in addition to statistical testing to better assess the goodness of fit of the binary logit models. To measure the model discrimination capability, the ROC curves for the models were reviewed. For the signalized intersection model, the ROC curve is generally smooth, showing model stability (see Figure 4-8, page 127). The unsignalized model produced a similar curve (see Figure 4-9, page 127). Finally, the pure segment model ROC curve, is the best of the three, demonstrating better discrimination for lower specificity values (see Figure 4-10, page 127). The pure segment model was found to have excellent discrimination with a c value of 0.82. All the competing models c values (area under the ROC curve) were above 0.7, which amounts to acceptable discrimination, following the guidelines set forth in Hosmer and Lemeshow (2000). Therefore the AIC criterion was the determinant factor for model goodness of fit comparison in this analysis.

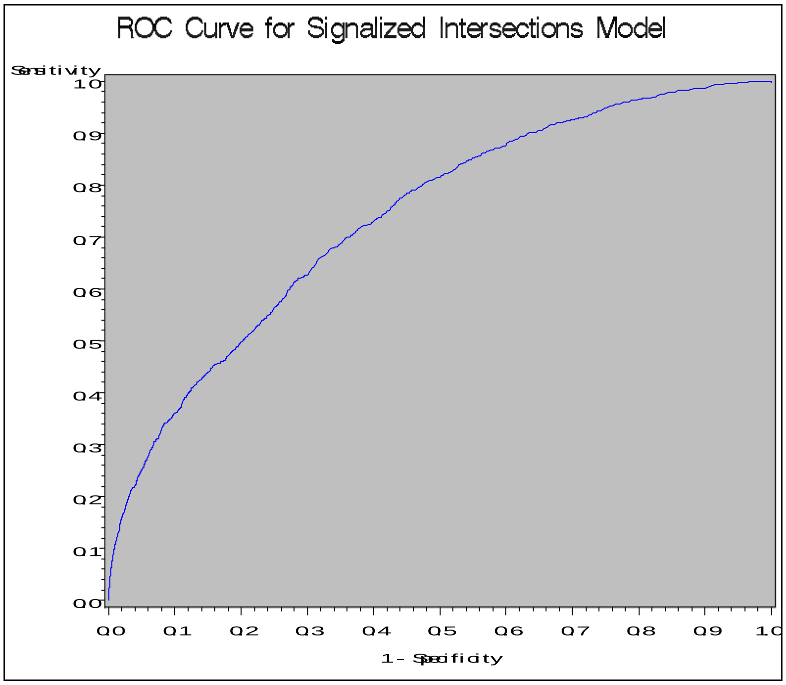


Figure 4-8: ROC Curve for the Signalized Intersection Crash Driver Injury Severity Model

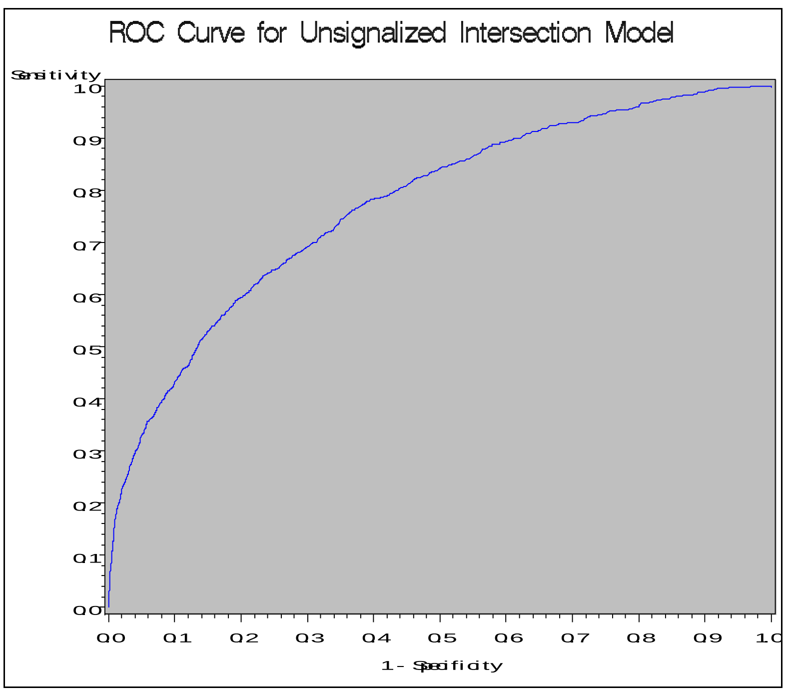


Figure 4-9: ROC Curve for the Unsignalized Intersection Crash Driver Injury Severity Model

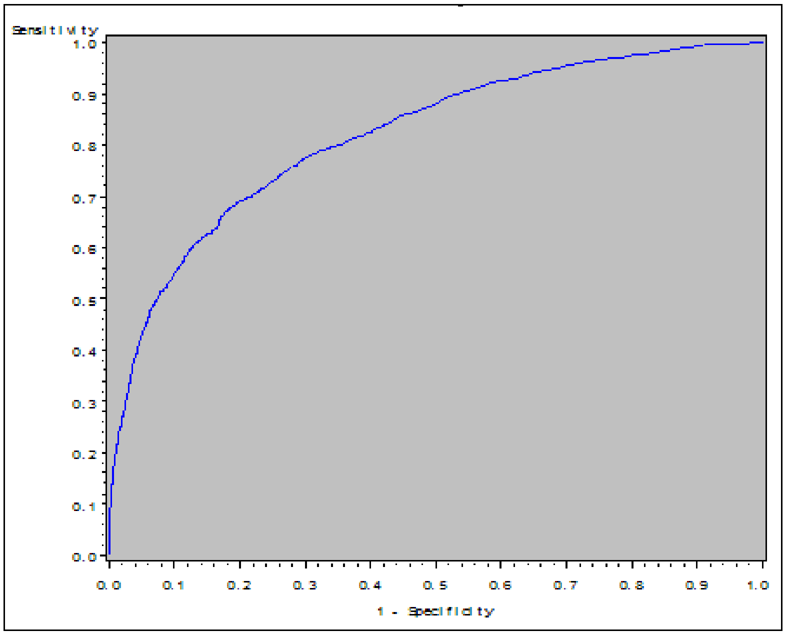


Figure 4-10: ROC Curve for the Pure Segment Crash Driver Injury Severity Model

The model building and assessment showed that the best results are obtained by classifying the intersection crash involvements by traffic control at the intersection of interest. Previous research has shown the validity of developing specific models for intersection crashes by traffic control and other geometric features. Some bias with respect to the total population may be present due to a larger proportion of signalized intersections on the state roads. However, the data availability benefit outweighs the possible bias, especially when considering multilane arterials, which tend to have similar characteristics over the entire road network. In summary, the resulting models' goodness of fit (refer to Table 4-22, page 126) shows that:

- The three best models by comparison of their AIC value are the unsignalized intersection, signalized intersection (multiple vehicle involvements), and pure segments.

- These three models have acceptable or excellent discrimination. In addition, the three models seems to fit quite well, as indicated by the p-value for the Hosmer-Lemeshow (>0.05), a test of model calibration. All models had previously passed the score and likelihood ratio tests (not shown here).

This analysis leads to the conclusion that, statistically, there is good reason to follow a modeling scheme of unsignalized intersection, signalized intersection (multiple vehicle involvements), and pure segments to appropriately explain the driver's injury severity for crash involvements on high-speed multilane arterials. However, these models failed to account for all the crashes because the single vehicle crashes at signalized intersections could not be included. In addition, other calibration and discrimination measures of fit were favorable for some of the aggregated models (i.e. all crashes and segments).

4.4 Conclusions from the Exploratory Analysis

In conclusion, different crash conditions and data availability influenced the results of this analysis. Data mining techniques and correlation analysis allowed for feasible first order interaction model building without sacrificing model stability. The most important variables for the crash involvements driver injury severity are: collision type (harmful event), safety equipment and ejection, gender, age, and estimated speed (in signalized intersections). Other moderately influential variables include driver fault, contributing cause (and aggressive driving), point of impact, speed ratio (for pure segments), alcohol/drugs (in unsignalized intersections), speed limit (in signalized and unsignalized intersections), and traffic volume (for pure segments). Regarding environmental factors, lighting condition is more influential in the unsignalized intersection model, while shoulder width is more so in the signalized model. The street lighting represents yet another design feature shown to reduce the driver injury severity in a crash involvement. These preliminary results will be extended and validated by multiyear crash analysis presented in Chapter 5.

There are three main models that were the best to identify the factors that contribute to crash involvement driver injury severity. However, a final conclusion could not be achieved because of three reasons. First, the data preparation process needs to be improved to avoid repeated values of crash, roadway and environment-related variables. Second, the model building and assessment process will be improved by better selecting categorical variable cut-off values and by expanding the preliminary analysis of each variable to improve pre-screening. Additional criteria will be examined for the model assessment in the final analysis. Third, the signalized intersection model must include single and multiple vehicle crashes, otherwise the models are not comparable since not all crashes are included and the modeling comparison may be misleading.

CHAPTER 5. FINAL ANALYSIS

5.1 Introduction

The main difference between the data in the exploratory analysis and the final analysis presented in this chapter is the format of the driver and vehicle data. While in the exploratory analysis all involvements were considered, in the final data preparation each crash had one row (record) which included up to four involvements. The main goal of the final analysis was to reach conclusions about the injury severity risk using logistic regression statistical models not affected by repeated measures. To achieve this, one involvement per crash can be analyzed in a model. We notice from Table 5-1 that most involvements (up to 82.6%) had two or less drivers involved. This means that involvements with driver sections 1 and 2 provide an appropriate sample to analyze the multiple vehicle crashes. The crash and involvement figures shown below were used in the preliminary analysis; the final analysis will be discussed in Section 5.4.

Table 5-1: Drivers Involved in Crashes in High-speed Multilane Roads

All involvements (state roads)		
Total Number of Drivers	**Frequency**	**Percent**
1	18701	8.66%
2	159912	74.07%
3	29337	13.59%
4	6362	2.95%
5	1252	0.58%
6	237	0.11%
7	64	0.03%
8	21	0.01%
9	7	0.00%
10	2	0.00%
11	2	0.00%
12	1	0.00%
Total	**215898**	**100.00%**

Complete data subset (state roads)		
Total Number of Drivers	**Frequency**	**Percent**
1	9238	7.67%
2	88983	73.89%
3	17332	14.39%
4	3893	3.23%
5	774	0.64%
6	134	0.11%
7	43	0.04%
8	16	0.01%
9	4	0.00%
10	2	0.00%
11	1	0.00%
12	1	0.00%
Total	**120421**	**100.00%**

Those reports without missing data in the fields (variables) considered for the final analysis shown in Appendix B are considered complete reports. The differences in proportions of crashes reported in vehicle sections 1 and 2 were very small (refer to Table 5-1, page 131). There is no evidence of a bias in the proportion of multiple vehicle crashes when the dataset is reduced to those crashes with complete reports. Previous studies selected single vehicle crashes or multiple vehicle crashes with only 2 involvements including Duncan et al. (1998), Kockelman and Kweon (2002), Toy and Hammitt (2003) as well as Ulfarsson and Mannering (2004). This is considered a good practice for a systematic analysis because a small group of multiple vehicle crashes (with more than 2 involvements) which represent less than 5% of the total crashes may not be well represented in the disaggregate injury severity analysis.

Another option considered in this investigation was to consider two models that include single and multiple vehicle crashes. To test if the order of the driver-vehicle sections in the two-vehicle crashes were statistically independent, the values in Driver 1 and Driver 2 sections were analyzed for each variable. See Section 3.1.1 for more information on the vehicle-driver sections in the crash report. The preliminary analysis presented in Section 5.3 showed significant dependence between the Driver 1 and Driver 2 sections for most of these variables. Therefore, a stratified random sample of two-vehicle crashes was selected for the multiple vehicle crashes. The single vehicle crashes on the complete dataset were added to complete the dataset for the final analysis. The final analysis research steps outlined in Section 1.3 are presented in subsequent sections of this chapter. Recall that only crash involvements on high-speed (40 mph or higher speed limit) multilane (4 or more lanes) roads were considered in this investigation.

5.2 Preliminary Analysis

5.2.1 Categorical Variable Analysis

The exploratory analysis demonstrated that one of the most pressing issues confronted was the selection of cut-off points and base levels for the categorical variables. Several techniques including categorical data analysis, optimized bins, and exploratory regression analysis were used. The variable categories were chosen to allow effective comparison when using 26 different data subsets. This meant that the variable cut-off points remained uniform during the analysis, while maintaining model stability in each case. By analyzing the different crash types in the preliminary analysis, some of the variables exhibited low cell values and numerical problems. This was tested thoroughly in the final analysis to allow a comparison of the effects found significant in each of the models. We will consider the driver, vehicle, road, crash and environment-related variables separately to facilitate the discussion of the results. Recall that a severe involvement was defined as one with an incapacitating or fatal injury outcome.

Table 5-2: Drivers Involved (in Sections 1 and 2) in Crashes at High-speed Multilane Roads

Driver-vehicle section 1

Year	Severe_driver1		Total
Frequency (Percent)	Non-severe	Severe	
2002	42827 (31.58)	2994 (2.21)	45821 (33.79)
2003	41247 (30.41)	2772 (2.04)	44019 (32.46)
2004	42966 (31.68)	2818 (2.08)	45784 (33.76)
Total	127040 (93.67)	8584 (6.33)	135624 (100.00)

Test of independence p-value=0.0589, Contingency coefficient= 0.0065

Driver-vehicle section 2

Year	Severe_driver2		Total
Frequency (Percent)	Non-severe	Severe	
2002	45089 (32.14)	2723 (1.94)	47812 (34.08)
2003	42628 (30.38)	2665 (1.90)	45293 (32.28)
2004	44648 (31.82)	2548 (1.82)	47196 (33.64)
Total	132365 (94.34)	7936 (5.66)	140301 (100.00)

Test of independence p-value=0.0055, Contingency coefficient= 0.0086

The possible temporal variation for three years of crash data was addressed by testing the statistical independence between the response variable (severe driver injury) and the year of the crash involvement. The analysis in Table 5-2, page 133, shows that there is a drop in the counts of crashes for the year 2003, which follows the general trend seen in the total crash counts listed in the Traffic Crash Statistics publication (FDHSMV, 2007). Furthermore, these results generally coincide with the proportion of severe crashes in the high-speed multilane arterials found in the exploratory analysis. The chi-square test of independence resulted in a p-value slightly greater than 0.05, so that the null hypothesis of statistical independence is not rejected for the driver 1 section. However, for the driver 2 section, the p-value is less than 0.05, rejecting the statistical independence hypothesis. This is a cause of concern if we are to select a dataset for the models based on the driver section in the crash report.

Table 5-3: Crash Severity Proportions for High-speed Multilane Roads (All Jurisdictions vs. State Roads)

Crash Injury Severity	All high-speed multilane roads		State high-speed multilane roads	
	Frequency	Percent	Frequency	Percent
Unknown	1198	0.55	788	0.5
No injury (PDO)	65009	30.11	47897	30.57
Possible Injury	68410	31.69	49963	31.89
Non-incapacitating Evident injury	53665	24.86	38551	24.6
Incapacitating Injury	24632	11.41	17183	10.97
Fatal Injury	2984	1.38	2306	1.47
Totals	215898	100	156688	100

A second issue addressed during the preliminary analysis was whether crashes reported in long forms that occurred in non-state roads would be significantly different than those occurring on state roads. The dataset of 215,898 crashes included crashes in state and non-state roads without eliminating any crash report available. Due to the data requirements explained in Section 3.2, only crashes on state roads will be considered in the regression analysis. The crash severity (highest injury severity in each crash) was analyzed for both the total dataset and the subset of state road crashes. The total dataset had 215,898 crashes (refer to Table 5-3); crashes in state roads totaled 156,688 records. There is a significant amount of crashes occurring on non-state roads which are relevant to the analysis of the safety performance of the high-speed multilane arterials. The proportions were found to be almost equal for the two datasets. This preliminary analysis suggests that there are no significant differences between the severe crashes in state vs. non-state roads for the high-speed multilane arterials. Further research should be developed when additional non-state road data become widely available.

The categorical analysis described in Section 3.3 was developed for a group of 47 variables that were extracted from the dataset prepared for the final analysis. Appendix B shows the results of the categorical data analysis for the entire dataset of crashes on all high-speed multilane arterials and a second analysis using only crashes occurring on a state road (also high-speed multilane arterial). This analysis did not show discrepancies in the statistical independence tests. Also, there were no major discrepancies in the measures of association (Contingency coefficient and Cramer's V) between each independent variable and the driver injury severity response variable. The comparative analysis between the crashes occurring on state and non-state roads showed no evidence of a significant difference between the two groups in regards to the injury severity. Additional validation might prove the transferability of the results presented in this chapter to non-state high-speed multilane arterials. On the other hand, further analysis of the measures of association comparing their respective results for the driver 1 and 2 sections did show some discrepancies, which will be discussed next.

Table 5-4: Relative Strength of Association with Driver Injury Severity (from Drivers 1 and 2 Sections)

Degree of association	Driver-related	Vehicle-Crash-related	Environment- and Roadway-related
Strong	*Vehicle_Fault_Code1 (or 2)* *Speeding1 (or 2)* Alcohol_Drug_Use1 (or 2) First_Safety_Equipment1	On_Off_Roadway *Ejected1 (or 2)* *Location_on_Roadway1(or 2)*	*TIME_GROUP* Location_Type Trafficway_Character *Site_Location*
Moderate	*Driver_Ageg_Group1 (or 2)* *First_Contributing_Cause1 (or 2)*	*CRASH_LANE* *Total_Number_of_Drivers* Type_of_Vehicle1 (or 2) Point_of_Impact1 (or 2) First_Harmful_Event1 (or 2)	Rural_Urban Lighting_Condition *CRRATECD (Median plus land use)*
Weak	Red_light_running1 (or 2) Sex1 (or 2) *Physical_Defects1 (or 2)*	Vehicle_Movement1 (or 2) Vehicle_Use1 (or 2)	Divided_Undivided Highway Type_of_Shoulder NUM_LEGS Median_type TYPEPARK First_Traffic_Control *Number_of_Lanes*
Very weak	Race1 (or 2) Residence_Code1 (or 2)	Vehicle_Special Functions1 (or 2) First_Vehicle_Defect1 (or 2)	Weather Road_Surface Condition Road_Surface_Type

**Variables in italics represent weak of very weak associations with the driver 2 injury severity*

The relative strength of association shown in Table 5-4 compared the measures of association by degree of freedom in order to ascertain their relative strength. For example, a contingency coefficient of 0.08 or above (for degrees of freedom greater than 10, contingency coefficient values 0.15 and above) was considered strong relative to the results of rest of the variables. The strength of association shown in the exploratory analysis (see Table 4-5) did not agree with the analysis just presented. Some notable exceptions are lighting condition, driver at-fault and driver age group (which now have a strong association); vehicle movement, race, number of lanes and type of shoulder (which now have a weak association). Differences between the two analyses are expected since the final dataset have excluded all crashes on full access control roads.

A close examination of the variables that were found to have the most relative significance in the exploratory analysis helped to understand the nature of their relationship with the driver injury severity. This categorical variable analysis showed which direction (sign) to expect if these variables were found significant in the severity models. Also, it was part of the basis for testing interactions in the final analysis.

5.2.1.1 Driver-related Variables

The most important variables considered here are the driver injury severity, the driver's age and gender. The first is the target of our analysis, the latter are the most significant effects found in the literature. The injury severity levels by driver section (see Table 5-5) include all the crashes on state high-speed multilane roads. Sampling drivers from sections 1 and 2 resulted in a proportion of driver injury severity similar to that of the vehicle-driver sections 1 to 4. Records with missing and invalid data (from any of the variables) were removed. In the case of injury severity alone, removed records (injury severity levels 0 and 6) accounted for 12.15% of the total driver 1 records and 3.66% of the driver 2 records.

Table 5-5: Total Frequency of Involvements by Injury Severity for Different Driver Sections

Injury Severity Level	Driver sections 1-4		Driver 1		Driver 2	
	Frequency	Percent	Frequency	Percent	Frequency	Percent
0	37130	7.98%	26333	12.11%	7359	3.65%
1	254424	54.69%	118957	54.71%	104714	51.91%
2	90812	19.52%	33488	15.40%	49559	24.57%
3	57579	12.38%	26135	12.02%	28515	14.14%
4	23241	5.00%	11129	5.12%	11019	5.46%
5	1903	0.41%	1326	0.61%	543	0.27%
6	94	0.02%	81	0.04%	12	0.01%
Total	465183	100.00	217449	100.00	201721	100.00

The complete data showed in the right portion of Table 5-1 corresponded to 120,421 crashes during the years 2002-2004. Two separate datasets were created for the preliminary analysis. The first consisted of 120,421 involvements of the first driver in a multiple or single vehicle crash. The second dataset of driver involvements (not from single vehicle crashes) had more complete records (n_2=127,819) because it was sampled independently from the driver section 1 dataset for the preliminary analysis. Crashes with complete information in the driver 2 section were included in the second dataset, even when the driver 1 section was incomplete. Some trends of single and multiple crashes can be identified in the tables discussed below.

Table 5-6: Driver Age Group by Injury Severity for Vehicle-driver Sections 1 and 2

Driver-vehicle section 1					Driver-vehicle section 2				
Driver Age Group1	Severe driver1		Total	Pct	Driver Age Group2	Severe driver2		Total	Pct
Frequency (Row Pct)	Non-severe	Severe			Frequency (Row Pct)	Non-severe	Severe		
15-19 years	14239 (94.45)	837 (5.55)	15076	12.52%	15-19 years	10476 (95.11)	539 (4.89)	11015	8.62%
20-24 years	16428 (94.29)	995 (5.71)	17423	14.47%	20-24 years	15737 (94.84)	857 (5.16)	16594	12.98%
25-64 years	69189 (93.75)	4615 (6.25)	73804	61.29%	25-64 years	84216 (94.24)	5147 (5.76)	89363	69.91%
65-79 years	9236 (91.50)	858 (8.50)	10094	8.38%	65-79 years	8508 (93.78)	564 (6.22)	9072	7.10%
80-98 years	3608 (89.66)	416 (10.34)	4024	3.34%	80-98 years	1652 (93.07)	123 (6.93)	1775	1.39%
Total	112700	7721	120421	100.00%	Total	120589	7230	127819	100.00%
Percent	93.59%	6.41%			Percent	94.34%	5.66%		

Test of independence p-value=<.0001, Contingency Coefficient=0.042

Test of independence p-value=<.0001, Contingency Coefficient=0.0158

The age of the driver was a contributing factor in the exploratory models presented in Section 4.2. Middle age drivers (25-64 years old) compose the majority of the involved drivers in both sections, driver section 2 (69.91%) somewhat higher than section 1 (61.29%). The major difference in the proportions is for the very young drivers (15-19 years old), which represent 12.52% of the involved drivers in section 1 vs. only 8.62% in section 2. This suggests an increased involvement of youngsters in single vehicle crashes. Clearly, their proportions of severe injuries are lower than for any of the age groups, as their physical condition is generally most favorable in case of a crash. For high-speed multilane arterials, there is a significant difference in the proportion of severe crashes for older drivers (refer to Table 5-6). When comparing the driver 1 and driver 2 sections, the effect of single vehicle (off-road) crashes is perceived as increasing the chance of severe crashes. This was investigated in the final analysis models by testing the interaction variable driver age and off-roadway crash.

Land use (as a surrogate of travel choice) may also influence the severe crash outcomes for different driver age groups, an interaction (age group and rural/urban) was also tested in the final analysis. Another set of interacting variables (driver age and driver at-fault) was tested to determine if the decreased severity odds ratio found in the exploratory analysis holds for all age groups, but it caused numerical problems (quasi-separation) as explained in Section 5.4.2. This result tends to confirm the theory of the driver at-fault bias in the driver injury severity modeling.

Table 5-7: Driver Gender by Injury Severity for Vehicle-driver Sections 1 and 2

Driver-vehicle section 1

Gender1	Severe driver1		Total	Pct
Frequency (Row Pct)	Non-severe	Severe		
Male	68800 (93.80)	4545 (6.20)	73345	60.91%
Female	43900 (93.25)	3176 (6.75)	47076	39.09%
Total	112700	7721	120421	100.00%
Percent	93.59%	6.41%		

Test of independence p-value=0.0001, Contingency Coefficient=0.011

Driver-vehicle section 2

Gender2	Severe driver2		Total	Pct
Frequency (Row Pct)	Non-severe	Severe		
Male	68662 (94.64)	3886 (5.36)	72548	56.76%
Female	51927 (93.95)	3344 (6.05)	55271	43.24%
Total	120589	7230	127819	100.00%
Percent	94.34%	5.66%		

Test of independence p-value=<.0001, Contingency Coefficient=0.0149

Gender was a variable of major importance in all of the exploratory models, as shown in Section 4.2.2. Different driver behavior and physiological characteristics play a role in the different outcomes of a crash event. It can be observed that females have a larger proportion of severe injuries, while males have a larger number of severe injury involvements (refer to Table 5-7, page 139). There seems to be an overrepresentation of male total and female severe involvements, when compared to the general population. However, there is no direct gender exposure measure of the driving population on arterial corridors.

Table 5-8: Safety Equipment Used by Injury Severity for Vehicle-driver Sections 1 and 2

Driver-vehicle section 1

Safety Equipment1	Severe driver1		Total	Pct
Frequency (Row Pct)	Non-severe	Severe		
None	10532 (82.01)	2310 (17.99)	12842	10.66%
Seat belt / Child Seat	100331 (95.50)	4728 (4.50)	105059	87.24%
Other	1837 (72.90)	683 (27.10)	2520	2.09%
Total	112700	7721	120421	100.00%
Percent	93.59%	6.41%		

Test of independence p-value=<.0001, Contingency Coefficient=0.2054

Driver-vehicle section 2

Safety Equipment2	Severe driver2		Total	Pct
Frequency (Row Pct)	Non-severe	Severe		
None	6684 (85.20)	1161 (14.80)	7845	6.14%
Seat belt / Child Seat	112402 (95.33)	5509 (4.67)	117911	92.25%
Other	1503 (72.86)	560 (27.14)	2063	1.61%
Total	120589	7230	127819	100.00%
Percent	94.34%	5.66%		

Test of independence p-value=<.0001, Contingency Coefficient=0.1569

The use of seat belts has been an important factor in reduced injury severity odds ratio for the exploratory analysis. As shown in Table 5-8, the rate of severe injuries is much higher when no seat belt use is reported. However, previous studies such as the one by Richardson et al. (1996) have found seat belt use over reporting for the non-severe crashes to avoid traffic fines. For severe crashes, it is usually possible for the police officer (or EMT) to determine if the injured occupant was using a seat belt.

The rate of seat belt use for non-severe injuries is quite high (between 87 and 92% for driver 1 and 2, respectively), while for the severe injuries is between 61 and 76%. The difference in the rates for the severe and non-severe crashes is between 17 and 28%. The official rate of usage across Florida in 2004 was 76.3% (FDOT, 2008). Comparing the rate of seat belt use of the non-severe crashes with the average use in Florida reveals that the over reporting could be as high as 13% for driver 1 and 17% for driver 2. Previous studies have suggested various rates of over-reporting of seat belt usage in non-severe crashes. A study by Streff and Wagenaar (1989) compared self-reporting of seat belt use to observational surveys of the same population. The authors' best estimate was to discount self-reported rates by 12 percent. In a study of police-reported crash data in Hawaii, Li et al. (1999) found a 10% reduction in the reported seat-belt use rate when adjusting for over-reporting. In addition, Hawaii hospital data showed that physicians reported 63.59% seat belt usage, while police reports usage rate was 90.26%. These figures are somewhat similar to the percentages presented in Table 5-8, page 140, which suggests that there is an over-reporting of the seat belt usage in the Florida crash data.

Table 5-9: At-fault Driver by Injury Severity for Vehicle-driver Sections 1 and 2

Driver-vehicle section 1

At Fault driver1	Severe driver1		Total	Pct
Frequency (Row Pct)	Non-severe	Severe		
Not cited	35730 (90.55)	3731 (9.45)	39461	32.77%
Cited	76970 (95.07)	3990 (4.93)	80960	67.23%
Total	112700	7721	120421	100.00%
Percent	93.59%	6.41%		

Test of independence p-value=<.0001,
Contingency Coefficient=-0.0867

Driver-vehicle section 2

At Fault driver2	Severe driver2		Total	Pct
Frequency (Row Pct)	Non-severe	Severe		
Not cited	112578 (94.23)	6899 (5.77)	119477	93.47%
Cited	8011 (96.03)	331 (3.97)	8342	6.53%
Total	120589	7230	127819	100.00%
Percent	94.34%	5.66%		

Test of independence p-value=<.0001,
Contingency Coefficient=-0.0193

The driver involved in a crash cited for a moving violation is considered at-fault. This measure of legal responsibility also reflects unsafe driving behavior. The lower rates of severe driver injury for those found at-fault just shown in Table 5-9, page 141, agree with the results of the exploratory analysis. It is important to note that the drivers in section 1 have a stronger tendency to be at fault; however, there is no evidence of a systematic bias for multiple vehicle crashes, as shown in Section 5.3. On the other hand, a large proportion of driver 1 *innocent* drivers sustained severe injuries (9.45%), which is one of the highest proportions seen so far in this investigation. At issue is whether single vehicle crashes with severe crash outcomes are less likely to involve *unsafe* driver behavior.

A study of Central Florida signalized intersections by Abdel-Aty (2003) found a significant negative effect of drivers not at fault, possibly due to the driver at fault being the striking vehicle, which for angle and turning crashes is expected to experience a lower level of injury than that of the driver of the stricken vehicle. This factor was found significant in all the exploratory models, which may have broader implications for the high-speed multilane arterials.

The speeding and contributing cause variables discussed next exhibit a similar situation with significant differences between the relationships for drivers 1 and 2. It is possible that police officers would tend to record cited drivers first and this compounded with the single vehicle crashes contributes to the higher proportion of severe injuries for the driver 1 section. Throughout this analysis, there is evidence of important differences between the driver 1 and 2 sections and serve as investigative support for the sampling of drivers from sections 1 and 2 for the final analysis. Only one driver involvement per crash was analyzed, and stratified sampling proved to be a sound method, as discussed in Section 5.3. This categorical data analysis served to indicate the association of these variables with the driver injury severity and to point out some trends that were ultimately confirmed in the final analysis. The tests for independence had the same conclusion for both sections in most cases, as shown in Appendix B. Only in one case were the variable was clearly significant in this test for both sections were excluded from further analysis. Part of the validation of the sampling technique involves comparing the final models with the trends presented in this section to confirm whether the sample captured the effects from both driver sections. These sections represent a vast majority of the total driver involvements.

Table 5-10: Driver Speeding by Injury Severity for Vehicle-driver Sections 1 and 2

Driver-vehicle section 1

Speeding1	Severe driver1		Total	Pct
Frequency (Row Pct)	Non-severe	Severe		
No speeding	23823 (89.04)	2933 (10.96)	26756	22.22%
Speeding	76846 (95.00)	4043 (5.00)	80889	67.17%
Unknown	12031 (94.17)	745 (5.83)	12776	10.61%
Total	112700	7721	120421	100.00%
Percent	93.59%	6.41%		

Test of independence p-value=<.0001, Contingency Coefficient=0.0993

Driver-vehicle section 2

Speeding2	Severe driver2		Total	Pct
Frequency (Row Pct)	Non-severe	Severe		
No speeding	23425 (91.39)	2206 (8.61)	25631	20.05%
Speeding	53356 (94.55)	3074 (5.45)	56430	44.15%
Unknown	43808 (95.74)	1950 (4.26)	45758	35.80%
Total	120589	7230	127819	100.00%
Percent	94.34%	5.66%		

Test of independence p-value=<.0001, Contingency Coefficient=0.0678

Speeding is suspected to be a major factor in severe crashes. Table 5-10 indicates that the rate of severe crashes is less for those drivers found speeding than for those not speeding. This result is similar to the at-fault driver results shown previously. The driver with the speeding citation would be at-fault and since it is usually the striking vehicle, it would cause severe damage to the stricken vehicle (usually not speeding). An interaction variable of speeding and point of impact was tested in the models to prove whether this theory was true. In addition, the proportion of non-speeding drivers in section 1 with severe injuries is of great concern.

Since the speeding indicator is computed using the estimated speed reported by the police officer only for certain types of crashes, there are more missing data than for any other variable. It is more likely that the police officer reports an estimated speed for a severe crash requiring a thorough investigation. Thus, it was deemed pertinent to include this variable (with one level labeled unknown) for its perceived significance in severe crash outcomes. The bias implications of this variable will be further discussed in the forthcoming sections.

Table 5-11: Driver Ejection by Injury Severity for Vehicle-driver Sections 1 and 2

Driver-vehicle section 1

Ejected1	Severe driver1		Total	Pct
Frequency (Row Pct)	Non-severe	Severe		
No	110866 (94.54)	6408 (5.46)	117274	97.39%
Yes or partial	1834 (58.28)	1313 (41.72)	3147	2.61%
Total	112700	7721	120421	100.00%
Percent	93.59%	6.41%		

Test of independence p-value=<.0001, Contingency Coefficient=0.2361

Driver-vehicle section 2

Ejected2	Severe driver2		Total	Pct
Frequency (Row Pct)	Non-severe	Severe		
No	118833 (94.98)	6285 (5.02)	125118	97.89%
Yes or partial	1756 (65.01)	945 (34.99)	2701	2.11%
Total	120589	7230	127819	100.00%
Percent	94.34%	5.66%		

Test of independence p-value=<.0001, Contingency Coefficient=0.1866

Ejection of the driver was found to be the most important variable affecting injury severity in the exploratory models. However, this is considered a post-crash event and in the exploratory models showed a propensity to cause numerical (quasi-separation) problems in some cases. This factor has been extensively considered in injury severity analysis on the interest of predicting its occurrence or effect on injury severity. As shown in Table 5-11, almost half of those drivers ejected in section 1 suffered severe injury. Given that the steering wheel turns into a source of injury, drivers without seat belts are expected to sustain higher degrees of injury.

Driver ejection has been primarily associated as a consequence of seat belt non-usage. However, it is not an exclusive determinant of severe injury.

As suggested by the reduced severe injury percentage for the drivers in section 2, driver ejection has been considered as an important part of the sequence of events in a roadside crash. Possible interactions are tested in the final analysis to investigate how this outcome is related to some crash precursors.

Table 5-12: Driver Contributing Cause by Injury Severity for Vehicle-driver Sections 1 and 2

Driver-vehicle section 1

Contributing Cause1	Severe driver1		Total	Pct
Frequency (Row Pct)	Non-severe	Severe		
No improper driver action	15244 (96.06)	625 (3.94)	15869	13.18%
Aggressive driving	32847 (92.97)	2484 (7.03)	35331	29.34%
Alcohol / Drugs	1828 (89.83)	207 (10.17)	2035	1.69%
Other	62781 (93.44)	4405 (6.56)	67186	55.79%
Total	112700	7721	120421	100.00%
Percent	93.59%	6.41%		

Test of independence p-value=<.0001, Contingency Coefficient=0.0441

Driver-vehicle section 2

Contributing Cause2	Severe driver2		Total	Pct
Frequency (Row Pct)	Non-severe	Severe		
No improper driver action	104122 (94.35)	6241 (5.65)	110363	86.34%
Aggressive driving	4900 (94.69)	275 (5.31)	5175	4.05%
Alcohol / Drugs	183 (80.97)	43 (19.03)	226	0.18%
Other	11384 (94.43)	671 (5.57)	12055	9.43%
Total	120589	7230	127819	100.00%
Percent	94.34%	5.66%		

Test of independence p-value=<.0001, Contingency Coefficient=0.0245

This parameter represents the most common driver actions that contribute to a crash, as reported by the police officer. The major categories used in this investigation were aggressive driving and alcohol/drug influence. Other categories had sparse data and would not provide practical results in the models used in the final analysis. As shown in Table 5-12, aggressive driving is an important factor in severe crashes, especially for the driver 1 section; 18.45% of the severe injuries driver involvements involved aggressive driving. However, when we compare the 29.34% aggressive driving involvements in section 1 with 4.05% in driver section 2, there is a possible bias in driver section a towards innocent (no improper action) drivers. The actions constituting aggressive driving include speeding, failed to yield right-of-way, improper lane change, followed too closely, improper passing and disregarded other traffic control. The implications of the differences found between drivers sections 1 and 2 are further discussed in Section 5.3.

Table 5-13: Driver Physical Defects by Injury Severity for Vehicle-driver Sections 1 and 2

Driver-vehicle section 1

Physical Defects1	Severe driver1		Total	Pct
Frequency (Row Pct)	Non-severe	Severe		
No	109808 (93.75)	7322 (6.25)	117130	97.27%
Yes	2892 (87.88)	399 (12.12)	3291	2.73%
Total	112700	7721	120421	100.00%
Percent	93.59%	6.41%		

Test of independence p-value=<.0001, Contingency Coefficient=0.0391

Driver-vehicle section 2

Physical Defects2	Severe driver2		Total	Pct
Frequency (Row Pct)	Non-severe	Severe		
No	119465 (94.34)	7163 (5.66)	126628	99.07%
Yes	1124 (94.37)	67 (5.63)	1191	0.93%
Total	120589	7230	127819	100.00%
Percent	94.34%	5.66%		

Test of independence p-value=0.963, Contingency Coefficient=-0.0001

The physical condition of a driver is expected to be a contributing factor in crashes. It was tested in the final analysis to find out whether this uncommon occurrence does have a significant impact in driver injury severity. The proportion of severe driver injury severity for section 1 is more than 12%, which is the highest for any variable reviewed so far (see Table 5-13). Most of the physiological conditions listed as defects are related to sight, hearing and fatigue. However, in the preliminary analysis other conditions such as, seizure, epilepsy or blackout registered above 20% severe injuries, the largest contributor to severe injuries due to the physical defects.

5.2.1.2 Crash and Vehicle-related Variables

In this section, some of the important variables from the exploratory models are examined. The effect of single and multiple vehicle crash involvements in driver injury severity is examined. Also, the main crash types found on high-speed multilane arterials are discussed.

Table 5-14: Crash Harmful Event by Injury Severity for Vehicle-driver Sections 1 and 2

Driver-vehicle section 1

Harmful Event Group1	Severe driver1		Total	Pct
Frequency (Row Pct)	Non-severe	Severe		
Rear-End	46732 (97.07)	1410 (2.93)	48142	39.98%
Head-On	2120 (87.75)	296 (12.25)	2416	2.01%
Angle	23953 (91.56)	2209 (8.44)	26162	21.73%
Left Turn	13342 (92.14)	1138 (7.86)	14480	12.02%
Right Turn	1446 (96.85)	47 (3.15)	1493	1.24%
Sideswipe	5524 (96.29)	213 (3.71)	5737	4.76%
Fixed Object	4208 (84.45)	775 (15.55)	4983	4.14%
Other	15375 (90.40)	1633 (9.60)	17008	14.12%
Total	112700	7721	120421	100.00%
Percent	93.59%	6.41%		

Test of independence p-value=<.0001, Contingency Coefficient=0.1403

Driver-vehicle section 2

Harmful Event Group2	Severe driver2		Total	Pct
Frequency (Row Pct)	Non-severe	Severe		
Rear-End	28344 (95.55)	1320 (4.45)	29664	23.21%
Head-On	2175 (88.96)	270 (11.04)	2445	1.91%
Angle	16082 (91.15)	1561 (8.85)	17643	13.80%
Left Turn	4790 (91.94)	420 (8.06)	5210	4.08%
Right Turn	584 (96.69)	20 (3.31)	604	0.47%
Sideswipe	4308 (96.90)	138 (3.10)	4446	3.48%
Fixed Object	1726 (90.37)	184 (9.63)	1910	1.49%
Other	62580 (94.97)	3317 (5.03)	65897	51.55%
Total	120589	7230	127819	100.00%
Percent	94.34%	5.66%		

Test of independence p-value=<.0001, Contingency Coefficient=0.0775

As presented in Table 5-14, page 147, the main types of collisions reported on the crash data under study are rear-end, angle and left turn crashes. Angle crashes are the highest contributors to severe injuries, accounting for 25.21% of the total, followed by the rear-end crashes (18.25%) and left turn crashes (10.4%). The proportions of crash types are clearly different in sections 1 and 2, more so for the left turn crashes. The fixed object crashes are still an important contributor to 6.41% of the severe injuries. Fixed object total involvements are one-tenth of the rear-ends, yet their corresponding severe injuries involvements are half of those attributed to rear-end in section 1. The final analysis included injury severity models by the four most important crash types (angle, rear-end, left turn and fixed object) which account for 60.31% of the driver severe injury involvements.

Table 5-15: Crash Point of Impact by Injury Severity for Vehicle-driver Sections 1 and 2

Driver-vehicle section 1

Point impact1	Severe driver1		Total	Pct
Frequency (Row Pct)	Non-severe	Severe		
Not driver's side	102882 (94.14)	6403 (5.86)	109285	90.75%
Driver's side	9818 (88.16)	1318 (11.84)	11136	9.25%
Total	112700	7721	120421	100.00%
Percent	93.59%	6.41%		

Test of independence p-value=<.0001, Contingency Coefficient=0.0707

Driver-vehicle section 2

Point impact2	Severe driver2		Total	Pct
Frequency (Row Pct)	Non-severe	Severe		
Not driver's side	109198 (94.61)	6216 (5.39)	115414	90.29%
Driver's side	11391 (91.83)	1014 (8.17)	12405	9.71%
Total	120589	7230	127819	100.00%
Percent	94.34%	5.66%		

Test of independence p-value=<.0001, Contingency Coefficient=0.0357

The point of impact is a surrogate measure of the crash mechanism. During the exploratory analysis, the best variable setting was to compare driver side impacts to every other point of impact. Driver side impacts are more likely in turn (left or right) and angle crashes, thus these are expected to be associated with higher driver injury severity. Table 5-15 shows the effect of single vehicle crashes in the proportion of severe crashes. There are 30% more severe injuries for drivers in section 1 when compared to section 2, if the impact is on the driver side. When the impact is not on the driver side, the difference is much less. This suggests that certain crash configurations at intersections should be examined more closely. This will be discussed in Section 5.4, with examples of interactions tested in the final analysis.

Table 5-16: Crash Vehicle Maneuver by Injury Severity for Vehicle-driver Sections 1 and 2

Driver-vehicle section 1

Vehicle Maneuver1	Severe driver1		Total	Pct
Frequency (Row Pct)	Non-severe	Severe		
Straight Ahead	69870 (93.34)	4987 (6.66)	74857	62.16%
Slowing / Stopping	20988 (92.34)	1742 (7.66)	22730	18.88%
Left Turn	4290 (96.82)	141 (3.18)	4431	3.68%
Changing Lanes	5688 (94.58)	326 (5.42)	6014	4.99%
Other	11864 (95.76)	525 (4.24)	12389	10.29%
Total	112700	7721	120421	100.00%
Percent	93.59%	6.41%		

Test of independence p-value=<.0001, Contingency Coefficient=0.0457

Driver-vehicle section 2

Vehicle Maneuver2	Severe driver2		Total	Pct
Frequency (Row Pct)	Non-severe	Severe		
Straight Ahead	60663 (93.11)	4491 (6.89)	65154	50.97%
Slowing / Stopping	8908 (93.48)	621 (6.52)	9529	7.46%
Left Turn	2471 (97.32)	68 (2.68)	2539	1.99%
Changing Lanes	1255 (96.09)	51 (3.91)	1306	1.02%
Other	47292 (95.94)	1999 (4.06)	49291	38.56%
Total	120589	7230	127819	100.00%
Percent	94.34%	5.66%		

Test of independence p-value=<.0001, Contingency Coefficient=0.0616

The vehicle maneuver helps explain the crash mechanism. The settings for this variable were changed slightly after analyzing some early results of the final analysis. One of the situations shown in Table 5-16 is that the left turns and changing lanes, which were used in the exploratory analysis, did not have enough involvements for regression modeling (by crash type). It was decided to combine the turning left, right and U-turn maneuvers (angle), as well as slowing/stopped and backing maneuvers (rear-end) due to their similar crash mechanism. The changing lanes maneuver was included in the others category. These results also show similar proportions of severe driver injury associated with the straight ahead and the slowing/stopping maneuvers. This may be due to the dominance of rear-end crashes and operating speeds lower than freeways, resulting in similar injury outcomes for the striking vehicle and the struck vehicle drivers. A lower proportion of drivers making the left turn (3.18% in section1, 2.68% in section 2) experienced severe injuries when compared to the drivers involved in left turn crashes (7.86% and 8.06% of drivers in sections 1 and 2, respectively). This suggests that in general drivers making a left turn are less likely to sustain severe injury than those hitting their vehicle.

Table 5-17: Private Vehicle Use by Injury Severity for Vehicle-driver Sections 1 and 2

Driver-vehicle section 1

Private vehicle use1	Severe driver1		Total	Pct
Frequency (Row Pct)	Non-severe	Severe		
No	7401 (97.09)	222 (2.91)	7623	6.33%
Yes	105299 (93.35)	7499 (6.65)	112798	93.67%
Total	112700	7721	120421	100.00%
Percent	93.59%	6.41%		

Test of independence p-value=<.0001, Contingency Coefficient=0.0371

Driver-vehicle section 2

Private vehicle use2	Severe driver2		Total	Pct
Frequency (Row Pct)	Non-severe	Severe		
No	8320 (96.85)	271 (3.15)	8591	6.72%
Yes	112269 (94.16)	6959 (5.84)	119228	93.28%
Total	120589	7230	127819	100.00%
Percent	94.34%	5.66%		

Test of independence p-value=<.0001, Contingency Coefficient=0.0291

The relationship between vehicle use and severe crash injury may help certain countermeasures. Even tough it was not found significant in the exploratory models; it is the interest of this investigation to find if there is a relationship between this variable and the average truck volume factor. Also, another concern is how drivers in private vehicle found at-fault are associated with severe crash injury. These are investigated in the final analysis in Section 5.4. It is important to determine if the commercial transportation or freight has an important effect on severe driver injuries on high-speed multilane roads. From the information in Table 5-17, page 150, the percentages of non-private vehicle drivers with severe crashes are significantly less than those in private vehicles. It is yet to be determined if those crashes involving trucks are a major cause of severe injuries.

5.2.1.3 Roadway and Environment-related Variables

Some roadway and environmental characteristics were found significant in the exploratory models. In that analysis, the roadway characteristics had more relative significance than the environment-related variables. The road surface condition was found significant, but not the weather variable. This suggests that the severe injury outcome is affected to a degree by roadway characteristics that could be improved by engineers in the design, construction and maintenance of high-speed multilane arterial corridors. On the other hand, weather-related variables are indirectly related to the road surface conditions, the friction course and skid resistance. This interaction is appreciated in the wet pavement crash analysis discussed ahead.

Table 5-18: Road Speed Limit by Injury Severity for Vehicle-driver Sections 1 and 2

Driver-vehicle section 1

Speed limit1	Severe driver1		Total	Pct
Frequency (Row Pct)	Non-severe	Severe		
Less than 40 mph	6133 (90.62)	635 (9.38)	6768	5.62%
40-45 mph	88378 (94.88)	4766 (5.12)	93144	77.35%
50-55 mph	15193 (90.17)	1657 (9.83)	16850	13.99%
60-70 mph	2996 (81.88)	663 (18.12)	3659	3.04%
Total	112700	7721	120421	100.00%
Percent	93.59%	6.41%		

Test of independence p-value=<.0001, Contingency Coefficient=0.1118

Driver-vehicle section 2

Speed limit2	Severe driver2		Total	Pct
Frequency (Row Pct)	Non-severe	Severe		
Less than 40 mph	4436 (90.49)	466 (9.51)	4902	3.84%
40-45 mph	97345 (95.08)	5036 (4.92)	102381	80.10%
50-55 mph	16238 (92.13)	1387 (7.87)	17625	13.79%
60-70 mph	2570 (88.29)	341 (11.71)	2911	2.28%
Total	120589	7230	127819	100.00%
Percent	94.34%	5.66%		

Test of independence p-value=<.0001, Contingency Coefficient=0.0685

Figure 5-1: Video Log Snapshot of an Arterial Corridor with 65 mph Speed Limit: State Road 10 in Gadsen County (RDWYID 50030000 Direction: East MP: 6.57)

The speed limit of a road is one of the most important design parameters and controls other aspects not seen in the crash information analyzed in this investigation. This analysis is limited to high-speed multilane arterials. However, the crashes at intersections involve other roads with speed limits different from those of the arterials. The multilane arterial corridors with speed limits of 40 and 45 mph carry a major proportion (65.56%) of the severe injuries (refer to Table 5-18, page 151). However, the proportions of severe injuries for roads with higher speeds range from 7.87% to 18.12%, which suggest an increased severe injury risk for crash involvements in higher speed (50-70 mph) roads. On the other hand the proportion of severe injuries for the minor roads (less than 40 mph) remains almost equal across driver sections

(9.38% and 9.51%). This shows an increased risk for the minor roads, but not as high as the one for the higher speed roads (especially 60-70 mph).

Another area of interest was the injury severity outcome for drivers entering the arterial corridor from other intersecting roads. Table 5-18, page 151, shows a considerable increase in the proportion of severe injuries for drivers on roads with higher speed limits (60-70 mph). These might be intersecting freeways (intersection-related crash) or 60 mph multilane arterials found in rural areas, such as a segment in State Road 10 in Gadsen County (refer to Figure 5-1, page 152) and State Road 25 in Alachua County (refer to Figure 3-1, page 54). Also, lower speed intersecting roads (usually minor roads) have a considerable proportion of severe driver injuries. A possible interaction between contributing cause and the speed limits was investigated in the final analysis to determine whether engineering or educational countermeasures would be more effective in these two cases, which amount to 14% of the total severe injuries in this analysis.

Table 5-19: Road Lighting Condition by Injury Severity for Vehicle-driver Sections 1 and 2

Driver-vehicle section 1

Lighting Condition	Severe driver1		Total	Pct
Frequency (Row Pct)	Non-severe	Severe		
Daylight / Dusk / Dawn	82768 (94.16)	5132 (5.84)	87900	72.99%
Dark with street lighting	24493 (93.48)	1707 (6.52)	26200	21.76%
Dark without street lighting	5439 (86.05)	882 (13.95)	6321	5.25%
Total	112700	7721	120421	100.00%
Percent	93.59%	6.41%		

Test of independence p-value=<.0001, Contingency Coefficient=0.0731

Driver-vehicle section 2

Lighting Condition	Severe driver2		Total	Pct
Frequency (Row Pct)	Non-severe	Severe		
Daylight / Dusk / Dawn	88696 (94.44)	5221 (5.56)	93917	73.48%
Dark with street lighting	26393 (94.80)	1449 (5.20)	27842	21.78%
Dark without street lighting	5500 (90.76)	560 (9.24)	6060	4.74%
Total	120589	7230	127819	100.00%
Percent	94.34%	5.66%		

Test of independence p-value=<.0001, Contingency Coefficient=0.0352

Changes in lighting conditions on high-speed multilane arterial corridors were a contributing factor with moderate to minor relevance in the exploratory analysis. Road lighting maintenance and its importance in preventing severe crashes at night has been a recent topic of discussion for the FDOT. High-speed multilane roads under state jurisdictions are more likely to have better lighting conditions, especially in urban areas. Better lighting conditions keep the drivers severe injuries in proportions of 6.52% (driver 1) and 5.20% (driver 2), comparable to daylight crashes for both driver sections (refer to Table 5-19, page 153). The prejudicial effects of lack of street lighting are evident. The proportion of severe crashes at night when there is no lighting (13.95% for driver 1 and 9.24% for driver 2) is almost double the daylight rates. The benefits of improvements in street lighting could potentially reduce up to 238 severe crashes each year, when holding all other conditions constant. This figure is just an incomplete estimate because it does not include the driver sections other than the first two, but it is nonetheless significant when considering the total costs of severe crashes. Additional data from RCI will better describe these benefits in the final analysis.

Table 5-20: Rural and Urban Land Use by Injury Severity for Vehicle-driver Sections 1 and 2

Driver-vehicle section 1

Rural/Urban	Severe_driver1		Total	Pct
Frequency (Row Pct)	Non-severe	Severe		
Rural	51848 (92.01)	4502 (7.99)	56350	46.79%
Urban	60852 (94.98)	3219 (5.02)	64071	53.21%
Total	112700	7721	120421	100.00%
Percent	93.59%	6.41%		

Test of independence p-value=<.0001, Contingency Coefficient=-0.0604

Driver-vehicle section 2

Rural/Urban	Severe driver2		Total	Pct
Frequency (Row Pct)	Non-severe	Severe		
Rural	55633 (93.40)	3930 (6.60)	59563	46.60%
Urban	64956 (95.17)	3300 (4.83)	68256	53.40%
Total	120589	7230	127819	100.00%
Percent	94.34%	5.66%		

Test of independence p-value=<.0001, Contingency Coefficient=-0.0381

The literature agrees on the importance of the land use in certain road and traffic conditions that affect crash occurrence and severity. In the exploratory models it was a significant contributing factor. Many design characteristics and traffic conditions depend on the land use and Table 5-20, page 154, shows that crashes occurring in rural areas account for 56.40% of the severe crashes in this analysis. Meanwhile more crash involvements were reported in urban areas, which account for 53% of the driver involvements. Of the crash involvements occurring on roads in rural areas, 7.99% (section 1) and 6.60% (section 2) result in severe injuries. The relative severe involvement ratio between the drivers involved in rural and urban crashes is 1.48 for all the involvements under study. Excluding other factors, there is 1.48 times the chance of a severe injury per driver involvement on roads in rural areas vs. each involvement on roads in urban areas. It could be inferred that a set of conditions in the rural areas contribute to a significantly higher rates of severe involvements. Other design characteristics will complement these results and point to effective countermeasures tailored to different land uses.

Table 5-21: Type of Shoulder by Injury Severity for Vehicle-driver Sections 1 and 2

Driver-vehicle section 1

Type of Shoulder	Severe driver1		Total	Pct
Frequency (Row Pct)	Non-severe	Severe		
Paved	26270 (93.65)	1780 (6.35)	28050	23.29%
Unpaved	28394 (91.88)	2508 (8.12)	30902	25.66%
Curb	58036 (94.42)	3433 (5.58)	61469	51.05%
Total	112700	7721	120421	100.00%
Percent	93.59%	6.41%		

Test of independence p-value=<.0001, Contingency Coefficient=0.0427

Driver-vehicle section 2

Type of Shoulder	Severe driver2		Total	Pct
Frequency (Row Pct)	Non-severe	Severe		
Paved	26812 (94.47)	1568 (5.53)	28380	22.20%
Unpaved	29603 (93.59)	2027 (6.41)	31630	24.75%
Curb	64174 (94.64)	3635 (5.36)	67809	53.05%
Total	120589	7230	127819	100.00%
Percent	94.34%	5.66%		

Test of independence p-value=<.0001, Contingency Coefficient=0.0189

Another road characteristic found significant in the exploratory models was the type of shoulder. Outside shoulders have been incorporated into road design mainly as a safety feature, but it has also proven to facilitate traffic flow as a rest area for incident management. Crashes on roads with curb shoulders report lower proportions of severe injuries, while roads with unpaved shoulders register the highest percentages (between 6.4 and 8.1 %) (refer to Table 5-21, page 155). The lack of paved shoulder is suspected to be a contributor to roadside single vehicle crashes, which tend to cause severe injury. Curb shoulders close to the edge of the traveled way are dangerous at high speeds. However, the presence of curb and gutter also indicate urban designs, which usually are better illuminated, carry more traffic and have lower operating speeds. These unobserved factors might be part of the perceived benefit of curbed shoulders. The shoulder width data may clarify the relationship between shoulders and driver injury severity.

Table 5-22: Road Surface Conditions by Injury Severity for Vehicle-driver Sections 1 and 2

Driver-vehicle section 1

Road Surface Condition	Severe driver1		Total	Pct
Frequency (Row Pct)	Non-severe	Severe		
Dry	94456 (93.39)	6687 (6.61)	101143	83.99%
Wet	17118 (94.64)	969 (5.36)	18087	15.02%
Slippery	749 (94.45)	44 (5.55)	793	0.66%
Icy	36 (87.80)	5 (12.20)	41	0.03%
Other	341 (95.52)	16 (4.48)	357	0.30%
Total	112700	7721	120421	100.00%
Percent	93.59%	6.41%		

Test of independence p-value=<.0001, Contingency Coefficient=0.0195

Driver-vehicle section 2

Road Surface Condition	Severe driver2		Total	Pct
Frequency (Row Pct)	Non-severe	Severe		
Dry	101900 (94.22)	6255 (5.78)	108155	84.62%
Wet	17614 (95.07)	913 (4.93)	18527	14.49%
Slippery	713 (94.56)	41 (5.44)	754	0.59%
Icy	38 (92.68)	3 (7.32)	41	0.03%
Other	324 (94.74)	18 (5.26)	342	0.27%
Total	120589	7230	127819	100.00%
Percent	94.34%	5.66%		

Test of independence p-value=0.0002, Contingency Coefficient=0.0131

The road surface condition reflects the prevailing weather conditions. Although the relative importance of this variable is lower than the rest, it might prove to be an important interaction term. Also, around the state of Florida severe weather events such as rainstorms are common. Weather conditions have been found significant in previous studies of injury severity. The proportions of severe injuries were not higher for adverse weather conditions than for dry conditions (refer to Table 5-22, page 156). The variable in the models combined the slippery or icy conditions to avoid sparse values.

Table 5-23: Traffic Control by Injury Severity for Vehicle-driver Sections 1 and 2

Driver-vehicle section 1

Traffic Control	Severe driver1		Total	Pct
Frequency (Row Pct)	Non-severe	Severe		
Other or none	57080 (93.13)	4208 (6.87)	61288	50.89%
Traffic signal or yield	46769 (94.75)	2592 (5.25)	49361	40.99%
Stop sign or flash lights	8851 (90.58)	921 (9.42)	9772	8.11%
Total	112700	7721	120421	100.00 %
Percent	93.59%	6.41%		

Test of independence p-value=<.0001, Contingency Coefficient=0.0481

Driver-vehicle section 2

Traffic Control	Severe driver2		Total	Pct
Frequency (Row Pct)	Non-severe	Severe		
Other or none	59707 (94.71)	3332 (5.29)	63039	49.32%
Traffic signal or yield	51084 (94.46)	2998 (5.54)	54082	42.31%
Stop sign or flash lights	9798 (91.59)	900 (8.41)	10698	8.37%
Total	120589	7230	127819	100.00%
Percent	94.34%	5.66%		

Test of independence p-value=<.0001, Contingency Coefficient=0.0364

Another important road variable is the traffic control at intersections. For high-speed multilane roads, the intersections with minor roads are generally controlled by either stop signs or traffic signals. Preliminary analysis in Section 4.1.3 showed a trend of increased severe injuries for the urban signalized intersections and rural unsignalized intersections. The statistics shown in Table 5-23 denote the trends of decreased proportions of severe injuries for the signalized intersections. Meanwhile, the stop controlled intersections exhibited significantly higher proportions of severe injuries (9.42% and 8.41% for sections 1 and 2, respectively). In addition, a smaller number of uncontrolled intersections are included with the road segments (without traffic control or speed limit control). These are captured in the intersection models and possible interactions in the final analysis. The comparison between the stop controlled and the no control yields a severe involvement ratio of 1.47. Excluding other factors, there is 1.47 times the chance of a severe injury per driver involvement at stop controlled intersections vs. each involvement at segments or uncontrolled intersections. The stop controlled involvements resulting in severe injuries represent 12% of the total severe injuries. Improvements at these intersections may have a large potential benefit.

After reviewing the exploratory analysis results, it was suspected that crashes reported as yield control in high-speed multilane roads were highly correlated to a traffic signal (right turn lane yield). In a sample from one year of crash data from five arterial corridors in two different counties; almost 50% of the crashes were in fact located at a right turn lane at a signalized intersection (see Table 5-24). The sample was deemed acceptable given the fact that yield control crashes represent less than 2% of the total crashes. In addition, flashing beacon control was found to have a correlation with the stop control crashes. Therefore, these two cases were grouped in the traffic control variable.

Table 5-24: Traffic Control Observed for a Sample of Crashes Recorded as Yield Control (N=49)

Type of traffic control	Percent
Traffic Signal	48.98%
Yield at ramp	12.24%
Yield at median	16.33%
Stop Control	4.08%
Yield in minor road	8.16%
Not applicable	10.20%
Total	100.00%

5.2.2 Continuous Variable Analysis and Variable Transformation

5.2.2.1 Continuous Variable Analysis

There was a group of continuous distributions for some variables that are related to the driver-, traffic-, and roadway-related factors that affect crashes and their outcomes. First, some descriptive statistics helped ascertain the scaling of these variables (see Table 5-25). The means, standard deviation and ranges of the variables are similar for both driver section 1 and 2. Histograms of these variables were also examined and no major differences were found between the distributions of drivers section 1 and 2. Since all of these variables are road- or traffic-related, no major differences are expected when comparing vehicle-driver sections.

Table 5-25: Descriptive Statistics for Continuous Variables

Driver-vehicle section 1						
Variable	**N**	**Mean**	**Std Dev**	**Sum**	**Minimum**	**Maximum**
***adt* per lane (in thousands)**	119946	8.11194	2.83967	972994	0.32	36.5
Median width (ft)	119946	25.2499	12.5586	3028624	2	148
Avg Truck Factor (%)	119946	5.94562	4.05988	713153	0	54.57
Skid Resistance Number	119946	36.82425	5.10355	4416921	2	68
Surface width (ft)	119946	29.99503	6.66249	3597784	18	60
Shoulder width (ft)	119946	3.48623	2.19078	418159	0	21

Driver-vehicle section 2						
Variable	**N**	**Mean**	**Std Dev**	**Sum**	**Minimum**	**Maximum**
***adt* per lane (in thousands)**	127345	8.21907	2.77156	1046658	0.32	36.5
Median width (ft)	127345	24.95359	12.23325	3177715	2	148
Avg Truck Factor (%)	127345	5.78039	3.72408	736104	0	54.57
Skid Resistance Number	127345	36.83053	5.0753	4690184	2	68
Surface width (ft)	127345	30.24613	6.6807	3851694	18	60
Shoulder width (ft)	127345	3.45205	2.18685	439601	0	21

A total of 949 records were eliminated from the sample due to data error or extreme values, as explained below. The *adt* per lane variable is the ratio of *adt* divided by the number of lanes of the main road section, divided by 1,000. The scaling of the variables was to avoid very high coefficient values in the regression models. Median width is defined by the RCI Field Manual as the distance (in ft) from inside edge of closest painted line (or *through* pavement edge) at one side of the median and measured straight across to the inside edge of the closest painted line (or *through* pavement edge) on the opposite side (FDOT, 2007). The media width distribution had an average of about 25 ft, which is a respectable figure in terms of road separation. Median width values above 150 ft were considered extreme and 294 records were excluded from the final sample. The average truck factor (AVGTFACT) is 5.94%, which indicates the important heavy truck activity taking place on Florida arterials that is present as a contributing factor in the preliminary models.

The main measurement of pavement friction used in the United States is the friction number (FN), also called skid number. Skid resistance is usually expressed as the static friction coefficient multiplied by 100 (1). The testing to obtain this measurement is usually performed using the locked wheel tester (ASTM E-274) which waters the pavement while locking a fifth wheel attached to a vehicle at the predetermined test speed, to simulate emergency braking. The skid resistance numbers (nSKTRESNM) are in a scale of 1 to 100 in the RCI data for the road section. The number available in the CAR database is the average along the section of road tested. The average skid number is above 35, which is deemed acceptable, as detailed in the next section.

According to the RCI Field Manual, the surface width (nSURWIDTH) on a divided highway is the pavement width between the edge of the inside through lane and the outer edge of through lane pavement at the outside shoulder (FDOT, 2007). The expected values for two lanes on each side of the road are of at least 18 ft. A total of 655 records with values less than 18 ft were discarded to avoid data error. The most common values observed during the graphical analysis were 24, 36 and 48 ft, which correspond to two, three, and four lanes of 12 ft on each side of the road.

Shoulder width (in ft) is measured from the outer edge of the outside lane stripe to the outer edge of the shoulder. The RCI shoulder data most widely available are for paved shoulders and the nSLDWIDTH data in CAR do not include unpaved shoulders. Although there was information for other kinds of shoulders as separate variables, the data were too sparse and sometimes conflicting and was not used. An average shoulder width of 3.5 ft is not the recommended for neither urban nor rural areas, but there is a large variability, as suggested by the 2.1 ft standard deviation. Inside shoulder information was very limited and not included in this analysis.

Table 5-26: Pearson Correlations and Independence Test for Continuous Variables

Pearson Correlation Coefficients Prob > |r| under H0: Rho=0		
	Severe_driver1 (n=119,946)	**Severe_driver2 (n=127,345)**
***adt* per lane (in thousands)**	-0.07205 <.0001	-0.04231 <.0001
Median width (ft)	0.06394 <.0001	0.02866 <.0001
Avg Truck Factor (%)	0.07333 <.0001	0.0312 <.0001
Skid Resistance Number	0.03207 <.0001	0.02045 <.0001
Surface width (ft)	-0.03835 <.0001	-0.01193 <.0001
Shoulder width (ft)	0.0364 <.0001	0.01658 <.0001

The Pearson correlations for these variables were tested with relation to the driver injury severity in each of the two sections under analysis. The results of the statistical tests are shown in Table 5-26, page 161. All the tested variables were found to have significant correlation to the driver injury severity variable, but when we compared the correlation coefficients, we can clearly see the strongest (and negative) correlation of *adt* per lane, which indicates that increased traffic volumes decreased the severe driver injuries. Smaller roads (nSURWIDTH-less lanes) indicated decreased numbers of severe injuries. On the other hand, median width had the strongest positive correlation, with increased median widths indicative of higher severe injuries. Likewise, larger values of the average truck factor, shoulder width and skid resistance number variables led to increased severe injury values. The implications of these effects are discussed in the model analysis section.

5.2.2.2 Continuous Variable Transformations

Some of the continuous variables were combined into groups or categories to better describe these factors in an injury severity model. Some of the issues that will be discussed include scaling, valid values (range), and justification for use of continuous distributions for some variables that could have been discrete. In addition, this section briefly outlines the reasoning for the categories selected for the models. Among the main methods used were: exploratory regression analysis, categorical data analysis, results from previous studies, design guidelines and standards. These methods were not used in an absolute fashion, rather a scientific process of inquiry through a series of steps led to an informed decision. Driver-, road- and environment-related continuous factors were transformed from continuous to categorical in order to enhance their interpretative power in the injury severity models.

The driver age variable has been briefly discussed previously. The choice of driver age groups was primarily based on the previous studies using Florida crash and driver data, such as those used by Abdel-Aty et al. (1998). The cutoff values for drivers as very young (15-19 years old), young (20-24 years old), middle (25-64 years old), old (65-79 years old), and very old (80-98 years). The use of 99 years was avoided due to the negligible frequency (less than 1%) and possible data error for unknown values.

The median width (median size) of the multilane roads should be as wide as practical to decrease the risk of head-on crashes and headlight glare at night. The Florida Greenbook standard calls for minimums (see Table 5-27). Based on these standards and after initial tests, the median widths equal of greater than 40 ft were tested against those less than 15 ft, between 15 and 19.5 ft, and between 19.5 and 40 ft.

Table 5-27: Minimum Median Width for Multilane Facilities (Source: Florida Greenbook, 2005)

Rural Highways

Design Speed (mph)	Minimum Width (ft)
55 and Over	40
Under 55	22

Urban Streets

Design Speed (mph)	Minimum Width (ft)
50	19.5
45 or Less	15.5
40 or Less **	10

** Paved medians used for two-way turn lanes or painted medians

The Skid Resistance number discussed previously was analyzed using different cutoff values based on the FDOT guidelines. Recall from Section 5.2.2.1 that skid resistance is expressed as the static friction coefficient multiplied by 100. Many highway agencies use the friction measurements for the purposes of rehabilitation, reconstruction and resurfacing of pavements. There are monitoring programs in place to take action at certain predetermined intervention levels. The FDOT has a skid hazard elimination program has a systematic skid test program which covers about 25-35 percent of the Interstate and Primary Systems per year, along with new pavements. Also, District Safety Engineers use a report of wet weather crashes and determine which sections of highway with 25 percent or more wet weather crashes need skid tests. Depending on the test results, the FDOT guidelines are used to determine whether the skid hazard warrants an improvement project, there is need for further review (for new pavements up to 18 months) or if the skid test indicates the pavement friction is acceptable. The FDOT skid resistance guidelines in the Skid Hazard Reporting System Manual are shown in Table 5-28 (FDOT, 2006).

Table 5-28: FDOT Friction Number Guidelines (Source: State Safety Office, 2006)

Posted Speed Limit (mph)	All highway sections surfaces		
	1 QUESTIONABLE	2 REVIEW	3 DESIRED
	FN40	FN40	FN40
Less than or equal to 45	25	26-28	30
Greater than 45	27	28-30	35

The values for the groups considered were analyzed and tested in preliminary regression models. Part of this analysis is shown in Figure 5-2 and Figure 5-3, both in page 165, which confirms the general increase in severe injuries as the skid number increases. Examining the rural area distribution, we find that there is an unexpected increase in the proportion of severe injuries after skid number 35. Meanwhile the urban area distribution behaves a lot closer to expectations, lowering the severe crash proportions after skid number 35. Additional implications of these findings are discussed in the final analysis section.

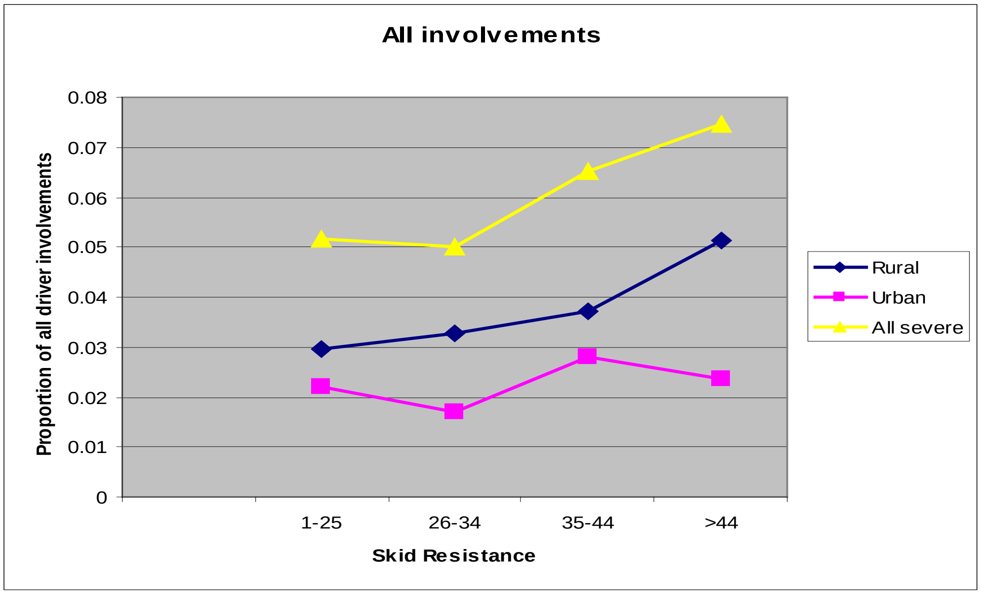


Figure 5-2: Distribution of Severe Injuries by Skid Resistance and Land Use for All Involvements

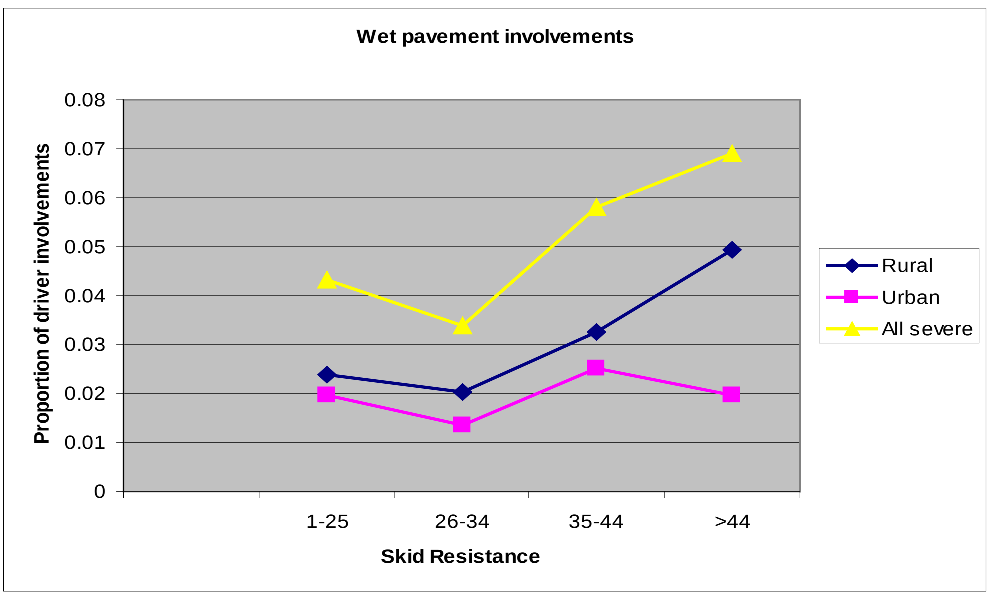


Figure 5-3: Distribution of Severe Injuries by Skid Resistance and Land Use for Wet Pavement Involvements

The analysis of Figure 5-2 and Figure 5-3, both in page 165, confirms that the general increase in severe injuries as the skid number increases for both the wet pavement crashes and the total involvements. The wet pavement severe injury to driver involvement ratio is generally lower than those for the total crash involvements. This comparison is necessary to compare the skid resistance to the type of crashes that it aims to reduce. The relationship seems to hold for both wet pavement and all involvements. The implications of the sudden increase after skid resistance number 44 are discussed later on.

The surface width (nSURWIDTH) variable was used in an interaction with number of lanes to derive a lane width variable. According to the RCI Field Handbook the surface width is measured across the traveled way, not including the shoulders (FDOT, 2008). The lane width was computed by dividing the surface width by half the number of lanes. The lane widths, as recommended by the Florida greenbook are 12 ft for major arterials and 11 ft for minor arterials. The minimum width in the standard is 10 ft; however, there is a significant number (8.59%) of driver crash involvements in roads with lanes less than 10 ft wide (see Table 5-29). This is a particular concern for the sections of road that are not complying with the current standards and their effect on the safety performance of important arterial corridors in the state.

Table 5-29: Driver Crash Involvements in High-speed Multilane Roads by Lane Width Group

Lane width group	Frequency	Percent
Lane width < 10 ft	9323	8.59
10 ft ≤ Lane width < 11 ft	6252	5.76
11 ft ≤ Lane width ≤ 12 ft	80889	74.56
Lane width > 12 ft	12030	11.09
Total	108494	100

The Florida greenbook recommends shoulders 10 ft wide in all roads. A minimum of 6 ft is required for roads with open drainage, while 8 ft is required for roads with heavy traffic volumes or a significant volume of truck traffic. The cutoff values for the shoulder width groups reflect this policy. An overwhelming majority (89.52%) of the driver crash involvements occurred on roads with shoulder less than 6 ft wide (see Table 5-30). Very similar proportions were found for the severe injuries. Another concern is that roads with larger shoulders have a little higher proportion of driver involvements. The wider road space has been associated with higher operating speeds, which in turn result in higher severe injury counts.

Table 5-30: Driver Crash Involvements in High-speed Multilane Roads by Shoulder Width Group

Shoulder width groups	Frequency	Percent
Shoulder width < 6 ft	97120	89.52
6 ft ≤ Shoulder width < 8 ft	3548	3.27
8 ft ≤ Shoulder width < 10 ft	3128	2.88
Shoulder width ≥ 10 ft	4698	4.33

The time variable was reduced to a binary variable denoting what is generally considered day and night hours after evaluation of exploratory analysis. There are some perceived negative effects of the involvements at night in severe injuries (see Figure 5-4, page 168). This variable did not show a strong correlation in the exploratory analysis or the preliminary analysis. The increased importance of the road characteristics vs. the environmental variables might be triggered by underlying correlations between weather characteristics and road characteristics. For example, road lighting is correlated to visibility at night and if this variable is significant denotes in part the effect of night crashes in the overall safety performance of high-speed multilane arterials.

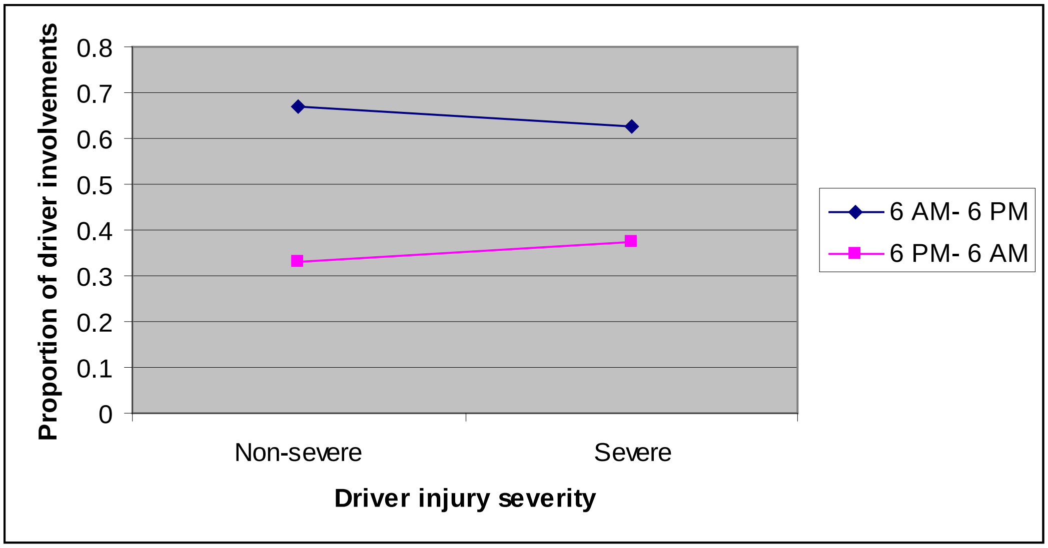


Figure 5-4: Distribution of Severe Injuries by Time of Day

5.3 Preliminary Regression Analysis

5.3.1 Severity Analysis by Road Entity

The driver injury severity distributions for each road entity selected for the analysis (see Section 3.6) was examined using the FDHSMV database (years 2002-2004) to compare the involvements occurring on high-speed multilane roads with all the involvements. The analysis in Figure 5-5, page 168, shows the totals for all of the involvements. The proportion of severe driver injury out of all driver involvements shows that in road segments, the likelihood of a severe injury given a crash involvement is the highest (6.38%), while the severe injuries at signalized intersections are the lowest (4.62%). The variability of the severe injury ratios is noticeable (38%) and thus the traditional analysis has separated both the crash frequency and injury severity models by road entity.

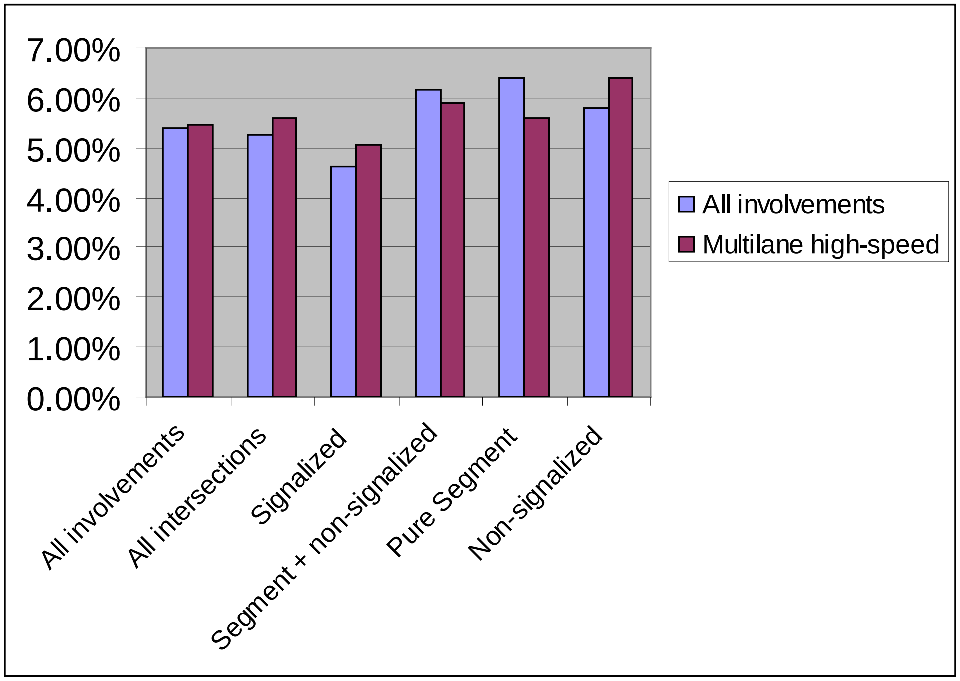


Figure 5-5: Distribution of Severe Injuries by Road Entities as a Proportion of Driver Involvements

On the other hand, the driver injuries at high-speed multilane roads exhibit a slightly higher severe ratio when compared to all of the involvements. The highest ratio of severe involvements occurred at unsignalized intersections, while the signalized intersections are the locations with lower ratios of severe driver injuries per involvement. This is different from the general perception of increased crash severity on curves and other road entities. This may be in part due to the less frequent roadway curves in the state arterials, when compared to other road types. The lowest injury to involvement ratio for the signalized intersections follows the general characteristic for all roads, but the ratio for high-speed multilane arterials is higher by almost 10% when compared to signalized intersections in all roads. In fact, all the road entities displayed higher ratios, with the exception of the pure segment and it’s derivate. These contrasts demonstrate that the high-speed multilane corridors have distinctive characteristics regarding their safety performance.

These results underline the major differences in the occurrence of crashes and injury severity outcomes at the high-speed multilane corridors. A difference is the increase of the severity ratios for intersections. When we compare the severe injury distribution in all roads to the distribution in high-speed multilane roads, the main difference is the variability. There is more parity between the involvements at (or related to) intersections and those occurring in road segments. When comparing all intersections and pure segments, the severity ratio is practically the same (5.58% vs. 5.59%). This suggests that a combined analysis of intersections and non-intersection is viable for involvements on high-speed multilane corridors.

5.3.2 Driver Involvements Selection for Analysis

When it is not possible to analyze all involvements for the injury severity analysis simultaneously, an appropriate sample must be selected. There are several alternatives to analyze driver involvements. One of the alternatives would be to select at-fault (or innocent) drivers as representative of the driving population. Another would be to select all driver involvements listed in section 1 (most used) of the crash report. A third alternative would be to select a sample of driver involvements from different sections, one for each crash. The first alternative was deemed inappropriate after confirming an association between driver at-fault status and the injury severity. The second alternative was first attempted using a sample of involvements from driver section 1 as a representative of all driver involvements because it includes all single and multiple vehicle crashes. To this end, the wide format discussed in Section 3.2.2 was prepared where driver sections would be arranged in one row per crash. Then, the complete records for each of the two driver sections were selected for analysis.

The results previously shown in Table 5-9, page 141, suggested a possible bias between driver sections and the at-fault driver status. Also, there is a concern that analysis using involvements from only one driver section will not be representative of the driver injury severity conditions of all crash involvements. First, the driver 1 and driver 2 variables were compared to find out if there was any apparent association between them. The last alternative was then tested to find out its effectiveness compared to the models already developed using a sample from one crash report section. The analysis previously shown in Table 5-5, page 137, indicated that a sampling process using driver sections 1 and 2 would be representative of the injury severity distribution of the total crash involvements. The discussion of the relationships between the driver 1 and 2 variables is presented in the next section.

5.3.3 Vehicle-driver Sections 1 and 2 Sampling Analysis

This analysis allowed a more detailed comparison between the variables vs. the driver 1 and driver 2 injury severities. Eleven out of the 21 variables are considered to have strong or moderate association with the driver 1 injury severity, but have a weak or very weak association with the driver 2 injury severities. Many of these variables (i.e. vehicle (driver) at-fault, speeding, ejected, driver age group) were found to have a significant effect on the injury severity in the exploratory analysis (see Section 4.2). This suggests a possible difference between the driver injury mechanisms of the two groups. This is expected, as the first section includes all the single vehicle crash involvements, whereas section 2 does not. A second important concern was shown in Table 5-2, page 133, where the proportions of severely injured drivers (especially in section 2) had statistical dependence with the year of statewide data selected. Minimizing temporal variation is very important to the validity of any further analysis.

Table 5-31: Driver Crash Involvements in High-speed Multilane Roads in the Complete Sample (All Involvements) and the Stratified Sample

Vehicle-driver Section	Stratified Sample		All involvements	
	Count	Percent	Count	Percent
Driver 1	55569	46.78	197197	50
Driver 2	63221	53.22	197197	50

An alternative analysis was performed by comparing the total crash involvements of drivers sections 1 and 2 and a stratified random sample of the driver 1 and 2 multiple crash involvements. The counts of the driver involvements in the datasets used are described in Table 5-31. The stratified sample consisted of 50% of the driver 1 and 2 involvements which had no missing data for the variables with significant association with the driver injury severity response (refer to Section 5.2). It was deemed important to keep a representative sample of the different driver involvements related to a single crash, while avoiding repeating crash data in the sampling process.

Using the function PROC SURVEYSELECT in a three-step process, a stratified random sample selection of the multiple vehicle crash reports was chosen. First, half of the involvements from section 1 were selected. Then, records with the report numbers in the selected driver 1 sample were deleted from the driver 2 dataset to avoid repeating crash data for the final analysis. Finally, a sample of 50% the original number of driver 2 involvements was selected. After selecting the sample of multiple crash involvements, single vehicle crash involvements were added to the sample. The final proportion of single and multiple crashes just shown in Table 5-31, page 172, were similar to those shown in Table 5-1, page 131.

Table 5-32: Driver Crash Involvements in High-speed Multilane Roads by Year and Injury Severity

Year	Driver Severity		Total
Frequency (Percent)	Non-severe	Severe	
2002	38232 (32.18)	2194 (1.85)	40426 (34.03)
2003	36434 (30.67)	2104 (1.77)	38538 (32.44)
2004	37779 (31.80)	2047 (1.72)	39826 (33.53)
Total	112445 (94.66)	6345 (5.34)	118790 (100.00)

Test of independence p-value= 0.0884

The analysis of the year to year variation of driver injury severity counts was also analyzed. The chi-square test of independence (see Table 5-32) resulted in a p-value greater than 0.05; thus the null hypothesis of statistical independence is not rejected for the entire sample. The p-value is higher when compared to the test show in Table 5-2, page 133, which suggests an improvement in the resistance to yearly variation for the stratified sample. This is a very important advantage of the stratified sample. Rather than magnifying the yearly differences, it becomes more heterogeneous in terms of the driver-, vehicle-, roadway- and environment-related characteristics that might be contributing factors to the driver injury severity.

Comparing the categorical analysis of the initial set of crashes and the final sample served to test the effects of drawing a sample of multiple crash involvements for the regression modeling. Additional analysis into the relationships between the driver 1 and 2 sections was performed by separating the single and multiple vehicle crashes. By comparing the counts of some of the most important variables, we can test whether the driver 1 and driver 2 sections are statistically independent (assigned randomly) or if there is any systematic relationship.

Table 5-33: Test of Independence between Driver Section Number and the Variables Listed (Sample n=118,790; Complete N= 394,394)

Variable	p-value using random sample	p-value using complete sample
Severe_driver_x	0.000298199	0.0002982
Year	0.000298199	1
Driver_Ageg_Group_x	0.177952724	<0.0001
Gender_x	<0.0001	<0.0001
Safety_Equipment_x	<0.0001	<0.0001
Speeding_x	<0.0001	<0.0001
Contributing_Cause_x	<0.0001	<0.0001
At_Fault_driver_x	<0.0001	<0.0001
Red_light_running_x	<0.0001	<0.0001
Residence_Code_x	<0.0001	0.00613153
Physical_Defects_x	0.325869667	<0.0001
Ejected_x	<0.0001	<0.0001
nRecommend_Re_Exam_x	<0.0001	<0.0001
nRace_x	<0.0001	<0.0001
Harmful_Event_Group_x	<0.0001	<0.0001
Off_Roadway	<0.0001	1
Point_Impact_x	<0.0001	<0.0001
Vehicle_Maneuver_x	<0.0001	<0.0001
Type_of_Vehicle_x	<0.0001	<0.0001
Private_vehicle_use_x	<0.0001	<0.0001
CRASH_LANE5	0.003356631	1
nRural_Urban	<0.0001	1
Location_Type	0.751214189	1
nVehicle_Special_Functions_x	0.00312239	<0.0001
nFirst_Vehicle_Defect_x	<0.0001	<0.0001
nCrash_Fault_Code	<0.0001	1
nTotal_Number_of_Drivers	<0.0001	1
nWork_Area_x	0.021756961	0.49171805
nAlcohol_Drug_Use_x	0.397420419	<0.0001

Partial results of the tests for each driver-, vehicle- and crash-related variable vs. driver section number are shown in Table 5-33. Complete results are shown in Appendix B. This analysis used multiple-vehicle crash involvements only. When using the random sample only the driver age, physical defects, location type and alcohol use variables that became significant for statistical dependence of the driver section. The complete results shown in Appendix B suggest that some possible numerical problems (quasi or complete separation) are possible if all involvements are used because of the repeated values of road characteristics. Some of these problems were apparent in the development of the exploratory regression models (see Section 4.2). The stratified sample showed a positive effect in these variables by alleviating the separation problems due to repeated values.

Table 5-34: Driver Crash Involvements in High-speed Multilane Roads in the Stratified Sample

Driver-vehicle section	Frequency	Percent
Driver 1 single vehicle	10587	8.18%
Driver 1 multiple vehicle	55591	42.96%
Driver 2 multiple vehicle	63235	48.86%
Total	129413	100.00%

Table 5-35: Goodness of fit for the Models using the Complete Records Driver 1 Section Dataset

GOF Parameter	OVERALL	INTERS	SIGNAL	SEGMENT	PURE SEG	UNSIG
Number of Variables	28	24	16	25	19	20
Degrees of freedom	52	41	31	46	35	38
Sample size	120442	70167	41779	71671	43283	28388
Response severe injury ratio	6.41%	6.23%	5.53%	7.21%	7.18%	7.25%
AIC	48211.51	28534.75	15861.77	30457.82	17747.43	12652.1
Hosmer-Lemeshow p-value	0.2124	0.2355	0.241	0.7759	0.2078	0.1405
c value (area under ROC curve)	0.789	0.764	0.759	0.803	0.824	0.771
Percent Concordant	78.5	75.9	75.3	79.9	82.1	76.7
Adjusted R-squared	0.1953	0.2355	0.1384	0.2229	0.2544	0.1835

Based on this analysis, it was decided to compare the road entity regression models using the two datasets to determine the best course of action for the final stage of this investigation. The composition of the database based on a sample of driver 1 and driver 2 sections is shown in Table 5-34, page 175, applicable to the six road entity models using driver section 1 using the complete records dataset. Table 5-35, page 175, shows the goodness of fit performance of these models. Although the goodness of fit measures for the first models was deemed acceptable, there was a comparison with models using the random sample of multiple vehicle driver involvements to assess not only the statistical model performance, but the coefficient interpretations, as suggested by Saccommano et al. (1994).

To make a comparison, models for the road entities using the stratified sample of driver involvements from sections 1 and 2 were developed. The goodness of fit for these models was also acceptable (see Table 5-36, page 177). This database was slightly reduced to 129,193 records for the regression analysis model due to invalid or missing data, including discarding crashes on road sections with very large medians (<150 ft), which did not change results significantly (not more than 3% of any odds ratio), but improved the median size coefficient. These few cases with very large median sizes might have been one-way roads or special cases which were not the main interest of this investigation. The Hosmer-Lemeshow p-values were improved, higher p-value is better in this calibration test. The coefficients values did not change drastically, but these models now take into account the diversity of driver involvements in multiple vehicle crashes to guard the efficiency of these models against certain biases, such as at-fault drivers, shown in previous sections.

Table 5-36: Goodness of fit for the Models using the Stratified Driver 1 and Driver 2 Records Dataset

GOF Parameter	OVERALL	INTERS	SIGNAL	SEGMENT	PURE SEG	UNSIG
Number of Variables	33	26	20	27	24	17
Degrees of freedom	67	56	43	60	48	48
Sample size	129193	73547	43944	77623	48020	29603
Response severe injury ratio	6.10%	6.01%	5.38%	6.75%	6.64%	6.94%
AIC	50600.81	29451.99	16614.84	31836.67	18897.62	12868.79
Hosmer-Lemeshow p-value	0.2493	0.5760	0.2507	0.8468	0.8790	0.1886
c value (area under ROC curve)	0.768	0.745	0.736	0.786	0.804	0.757
Percent Concordant	76.2	73.9	72.8	78.1	79.9	75.2
Adjusted R-squared	0.1801	0.1482	0.1236	0.2109	0.2388	0.1774

After evaluating and comparing these two sets of models, major differences were found. These results from these preliminary models were encouraging and a decision was made to proceed with final model development using the stratified sample. Some of the key advantages found are summarized next. First, the year to year statistical dependence of the driver injury severity was significantly reduced, which improves the validity of the analysis. Secondly, the ratio of severe injuries from the data using the stratified sample more closely resembles the total involvements on high-speed multilane arterials (see Table 5-37). The higher severe involvements response ratios are expected in the models because incomplete records were removed, which are likely minor crashes with no or lesser injuries.

Table 5-37: Severe Injury to Driver Involvement Ratios for Complete Driver 1 and Driver 2 Records

Road entity group	Non-severe	Severe	Total	Severe ratio
All involvements	275143	15847	290990	5.45%
All intersections	155727	9205	164932	5.58%
Signalized	95274	5065	100339	5.05%
Segment + Unsignalized	161781	10141	171922	5.90%
Pure Segment	101328	6001	107329	5.59%
Unsignalized	60453	4140	64593	6.41%

A third key advantage was that the random sample model generally had higher AIC values due to the increased heterogeneity of the data. However, this is usually a desired property on systematic crash analysis and statistical analysis in general. Minor loss of explanatory power (as measured by the adjusted R-squared value) was necessary to achieve more accuracy. There is an improved calibration of the models using the sampled multiple vehicle involvement data. Because the more homogeneous data do not completely reflect the variations in driver injury severity in multiple vehicle crashes, the statistical results might be misleading. Even tough the models suffer from a reduction in explanatory power, additional precision outweigh this loss.

A fourth advantage was the numerical stability of the random sample models was vastly improved compared to the earlier models. A set of covariates that more accurately represented the changes in driver injury severity was obtained. The coefficient significance showed a small improvement in the random sample models. Although a more heterogeneous dataset was used, the standard errors remained in the same order of magnitude.

Another advantage during model building was that the positive impact of interactions was noticed in the random sample models. Some important interactions in the driver 1 and 2 sections models did not significantly improve the model (AIC<10) and were eliminated. In the random sample models, important interactions were significant in the models, without adversely affecting significant main effects. During model building using the driver section data, interactions would cause dropping important main effects.

Finally, the variables found significant in the random sample models were more useful when compared to those in the earlier models. One of the reasons was the superior numerical stability of the random sample dataset. Important variables such as shoulder width and lane width were tested in both model sets; however, these were significant only in models using a sample of multiple vehicle involvements. Another important contribution was the addition of some roadway-related variables in the final stage of analysis, which is discussed in Section 5.4.1.

The evidence just presented and the variables that were found significant show that a random sample of multiple vehicle crashes from sections 1 and 2 plus the single vehicle crashes is more representative of the total driver involvements than involvements from one section only. Therefore, a sample of multiple and single vehicle crashes, one involvement per crash was selected for final analysis. A model with this kind of sample is expected to have better reliability and scientific validity.

5.4 Final Regression Analysis

5.4.1 Additional Roadway Data from RCI

Having selected an appropriate sample of driver involvements, the investigation focused on a preliminary analysis of additional roadway data made available before the final analysis. Initially, 15 road-related variables were available; these were reduced to 9 (see Table 5-39, page 181). From the original 129,413 crashes there was a loss of 20,897 crashes that were located on the borders of two road sections, as defined by RCI. The design of the RCI database follows a Linear Referencing System that joins each link to the next one by a common milepost number. It was found convenient to eliminate these crashes to avoid errors in assigning sections due to the large sample available for the final analysis. Additional invalid or missing values further reduced the dataset to 107,449 records for the final analysis (see Table 5-38, page 180); these were sufficient for the analysis. The number of vehicles involved remained similar to the original data contained in Table 5-1, page 131. In addition, the severity response ratios remained similar to those of the total involvements. There was no evidence of bias by eliminating these records.

Table 5-38: Count of Vehicles (Driver) Crash Involvements for Stratified Driver 1 and Driver 2 Records

Number of Vehicles	Frequency	Percent	Cumulative Frequency	Cumulative Percent
1	9354	8.71	9354	8.71
2	78230	72.81	87584	81.51
3	15510	14.43	103094	95.95
4	3485	3.24	106579	99.19
5	699	0.65	107278	99.84
6	119	0.11	107397	99.95
7	38	0.04	107435	99.99
8	8	0.01	107443	99.99
9	2	< 0.01	107445	99.996
10	2	< 0.01	107447	99.998
11	1	< 0.01	107448	99.999
12	1	< 0.01	107449	100.000

The new variables included in this section were important to the development of the final models because they captured *unobserved* information in the previous models. A categorical data analysis was performed for the new variables (raw data n=150,286) (see Table 5-39, page 181). Variables with too many missing values or non-significant independence tests were excluded from further analysis. The variables nACMANCLS, nAUXLNTYP and nAUXLNUM had missing values that could be included in the model because these indicated that these features were not present. The variables in bold were included in the final analysis: LIGHTCDE (non-high mast lighting pole density), LIGHTING (high mast lighting pole density), nACMANCLS (access management class), nAUXLNTYP (auxiliary lane type), nAUXLNUM (auxiliary lane number), nFRICTCSE (type of friction course), nPAVECOND (pavement condition), nSIDEWALK (sidewalk width) and, nURBSIZE (urban size).

Table 5-39: Test of Independence between Driver Injury Severity and the Variables Listed (n=150,286)

Variable	Missing	DF PCHI	chi-square	p-value	CONTGY	CRAMV
LIGHTCDE	0	6	78.77065	<0.0001	0.022888	0.022894
LIGHTING	794	4	13.54608	0.008894	0.009519	0.009519
ATTLOCCD	148698	5	6.086663	0.297877	0.061792	0.061911
nACMANCLS	23907	80	711.7266	<0.0001	0.074834	0.075045
nATTYPECD	148746	8	3.665354	0.885985	0.048728	0.048786
nAUXLNTYP	57619	19	31.64173	0.034283	0.018475	0.018479
nAUXLNUM	59511	9	16.2857	0.06115	0.013393	0.013394
nFRICTCSE	5282	10	83.68075	<0.0001	0.024016	0.024023
nPAVECOND	35	87	326.719	<0.0001	0.046581	0.046631
nRDSIDTYP	149444	6	19.97368	0.002799	0.152224	0.154019
nSIDEWALK	35	30	647.5108	<0.0001	0.065506	0.065647
nSIGNALNC	143512	2	32.87606	<0.0001	0.069497	0.069665
nSIGNALTY	74912	3	42.30534	<0.0001	0.023685	0.023691
year	0	2	1.484014	0.476157	0.003142	0.003142
nTURNMOVE	148333	4	10.48573	0.032994	0.073078	0.073274
nURBSIZE	462	14	1011.463	<0.0001	0.081889	0.082164

The variables related to road features included a categorical measure of street lighting (for high mast and non-high mast), the width of sidewalks, the type of auxiliary lane (if present) and the number of auxiliary lanes. Other variables related to the pavement included a condition index and the type of friction course included. This type of friction course is a feature independent from the measure of the friction (or skid resistance) number, but possibly related to friction performance in a way that could serve to improve the interpretative power of this model. Additional variables that further improved the information about land use were the urban size and access management class variables. The variables used in the final analysis are presented in detail in Appendix C.

5.4.1.1 Additional RCI Categorical Variable Setup

Many variables in RCI were already categorical in nature, other were transformed from numerical to categorical. A description and reasoning behind these categorizations is presented in this section (categorical variables) and the next section (numerical variable transformations). The lighting conditions on high-speed multilane corridors are further described by a categorical variable in RCI that describes the density of lighting poles in a road section. Both LIGHTCDE (non-high mast lighting pole density) and LIGHTING (high mast lighting pole density) have three categories: yes, partial and none. The categories are defined in Table 5-40, which indicate the different values for each category in high mast and non-high mast lighting. High mast lighting has various advantages including increased coverage area and better uniformity to avoid driver glare effects. In addition, high masts are generally relocated out of the clear zone and with significantly lower amounts per mile, may also contribute to improved roadside safety performance. In Florida, high mast lighting is primarily limited to interchange locations far from developed areas and some older sites due to light trespassing issues. The implications of these design differences will be discussed later in this section.

Table 5-40: Definitions of the Lighting Conditions of Roads from RCI

Codes	Non-high mast (LIGHTCDE)	High mast (LIGHTING)
N	One or none light poles exists in the section	One or none light poles exists in the section
P	Partial lighting exists (Rates of 4-24 lights poles per mile)	Partial lighting exists (Rates of 4-9 light poles per mile)
Y	Full lighting exists (Rates of 25 lights poles per mile or more)	Full lighting exists (Rates of 10 light poles per mile or more)

The Access Management Class variable (nACMANCLS) has a category 1 for limited access and classes 2 through 7 for the multilane non-limited access. Crashes on roads with category one (limited access) was not found in the final sample, as expected. The variable coding used in the final analysis follows the codes 2-7, as defined in Table 5-41, and one category for not applicable locations without an access classification by FDOT. Class codes 2-4 are defined by less dense land uses with intersection spacing of half a mile (lower density urban areas) and these categories were joined in the models, as the model building process dictated. In the end, the variable had five levels: class 2-4, class 5, 6, 7 and 9 (not applicable). The last level (9) was created for those road sections without access class codes.

Table 5-41: Definitions of the Access Class Codes (from Rule 14-97) of Multilane Roads in Florida (Source: FDOT, 2007)

<table>
<tr><th rowspan="2">Class</th><th rowspan="2">Medians</th><th colspan="2">Median Openings (ft)</th><th rowspan="2">Signal spacing (ft)</th></tr>
<tr><th>Full</th><th>Directional</th></tr>
<tr><td>1</td><td>Limited Access</td><td>N/A</td><td>N/A</td><td>N/A</td></tr>
<tr><td>2</td><td>Restrictive w/ Service Roads</td><td>2,640</td><td>1,320</td><td>2,640</td></tr>
<tr><td>3</td><td>Restrictive</td><td>2,640</td><td>1,320</td><td>2,640</td></tr>
<tr><td>4</td><td>Non-Restrictive</td><td></td><td></td><td>2,640</td></tr>
<tr><td>5
(>45 mph posted speed)</td><td rowspan="2">Restrictive</td><td>2,640</td><td rowspan="2">660</td><td>2,640</td></tr>
<tr><td>5
(≤45 mph posted speed)</td><td>1,320</td><td>1,320</td></tr>
<tr><td>6</td><td>Non-Restrictive</td><td></td><td></td><td>1,320</td></tr>
<tr><td>7</td><td>Both Median Types</td><td>660</td><td>330</td><td>1,320</td></tr>
</table>

No changes in the urban size (nURBSIZE) variable coding had five levels based on the area population: rural, small urban, small urbanized, large urbanized and metropolitan. The rural/urban binary variable in the model was substituted by urban size to investigate its effects on the model. The urban size did not perform as well as the binary rural/urban classification in the injury severity model. It could not be present simultaneously with the rural/urban variable due to collinearity issues (which were evident during model building), thus it was not entered into the final model.

Two variables related to the road geometric design were expected to contribute to improved intersection models. The auxiliary lane type (nAUXLNTYP) indicates the type of auxiliary lane (if present) in the road section or intersection where the crash occurred. The main types of auxiliary lanes tested were left turn, right turn, bus, merging (outside), merging (inside) and parking lanes. In the final models (segment and unsignalized) only the merging (outside), merging (inside) and parking lanes were significant. This variable is complemented by the auxiliary lane number (nAUXLNUM), which depicts the number of auxiliary lanes present (if applicable) in the road section or intersection where the crash occurred. The combination of these two variables was tested in all the models, with limited success.

Friction courses are applied to roads with heavy traffic volumes and high speed limits. The type of friction course has been evolving during the past decades and the field nFRICTCSE records eight types of friction courses, as defined by the FDOT. No crash data were found for the newer friction courses (FC-9.5 and 12.5). There are two general types of friction courses currently used by FDOT: dense graded and open graded. Their thickness is controlled by specification through the minimum and maximum spread rate. Generally friction course type 5 (FC-5) is specified for multilane roadways with speed limits greater than 45 mph. FC-6 mixes (dense graded) are typically specified for roadways with posted speed limits less than or equal to 45 mph, but were non-significant in the injury severity model essentially due to its low sample size. Older dense graded friction courses (FC-1 and FC-4) were found significant and are discussed in more detail in subsequent sections. The dominant type of friction course for the roads in which crashes were reported was FC-2 (see Table 5-42). Preliminary analysis and model building led to five categories: FC-2 (base), FC-1, FC-4, FC-5 and none or other types.

Table 5-42: Types of Friction Courses Related to Crash Involvements in High-speed Multilane Roads

Type Friction Course	Frequency	Percent
0 (none or null)	6767	6.30
1 (FC 1)	6736	6.27
2 (FC 2)	48354	45.00
3 (FC 3)	14022	13.05
4 (FC 4)	24618	22.91
5 (FC 5)	2885	2.68
6 (FC 6)	4067	3.79
Total	107449	100.00

An attempt was made to find correlations between friction course, skid resistance and severe injury to driver involvement ratios. Additional analysis of the interaction plot in Figure 5-6, page 186, shows some important correlations that should be taken into account when interpreting the results of the injury severity analysis models. First, older friction courses (FC-2 and FC-4) exhibited the highest severe injury ratio for skid resistance values of 35 and over. For road sections (or intersections) with FC-4, severe injuries account for 2.50% of total involvements for skid numbers over 44, compared to only 0.14% for all skid resistance values. For road sections (or intersections) with FC-2, severe injuries account for 3.18% of total involvements for skid numbers ranging from 35 to 44, but is also high (3.10%) for the rest of the skid resistance values.

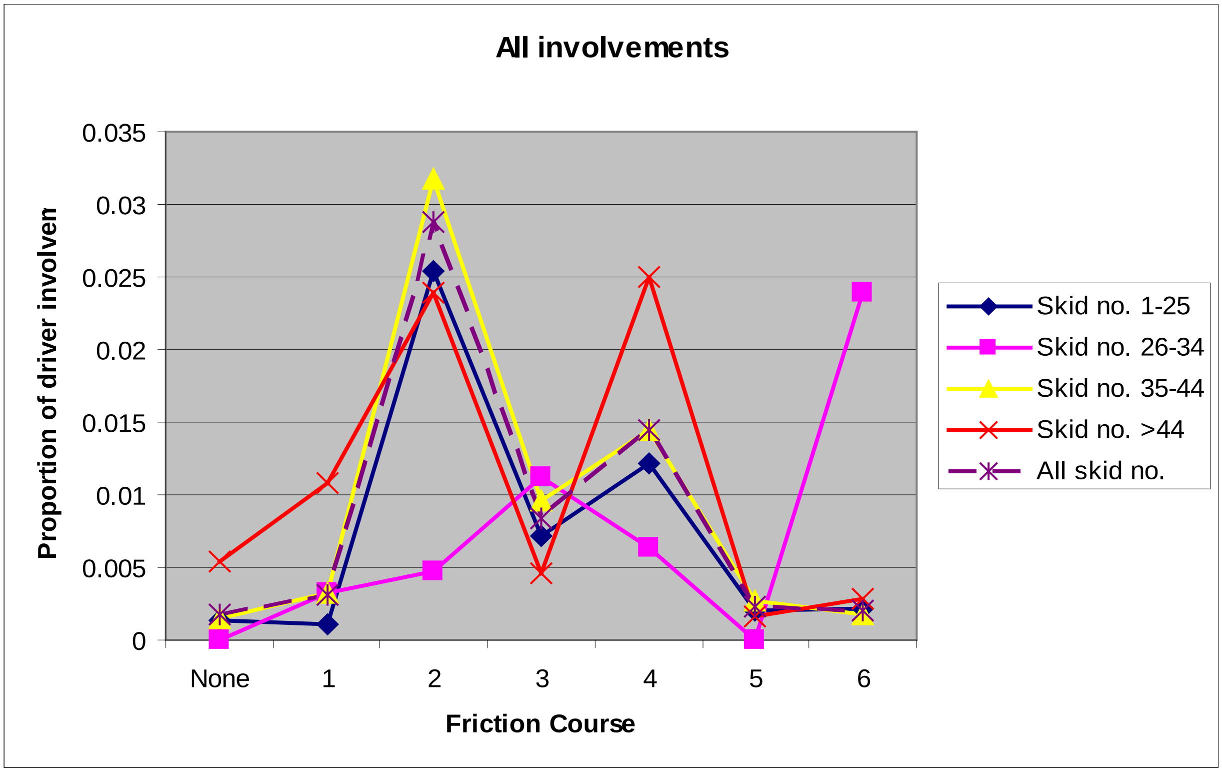


Figure 5-6: Severe Injury Ratio to All Involvements by Skid Resistance and Friction Course

Roads with friction courses FC-2 and FC-4 accounted for 46% and 23% of the total severe involvements. Similarly for the wet pavement crashes roads with FC-2 and FC-4 accounted for 50% 20% of the severe injury involvements, as shown in Figure 5-7. There is a clear tendency of increased total and severe injury involvements at locations with older friction courses. Decreasing skid resistances of older friction courses (polishing effects) is an important concern for skid hazard prevention programs. However, roads under wet pavement hazards are considered when at least 25% of the crashes are related to wet pavement. If there is a systematic decrease in friction resistance on high-speed multilane corridors with older friction courses, it is not necessarily captured at the district level. The injury severity models only showed a trend by land use.

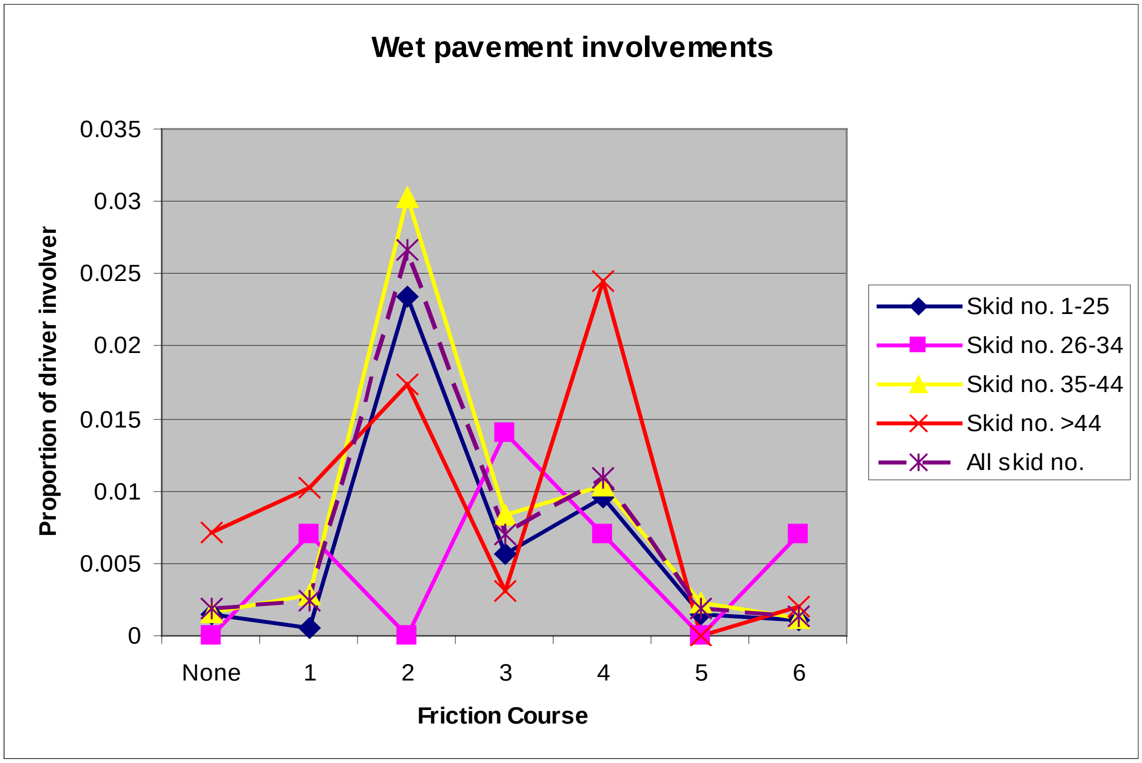


Figure 5-7: Severe Injury Ratio to Wet Pavement Involvements by Skid Resistance and Friction Course

Another important concern is the frequency of severe injuries on roads with skid numbers 1-25 (deemed unsafe) for both total and wet pavement crashes (see Table 5-43). Driver crash involvements on locations with low skid resistance accounted for 28.56% of the severe injuries (total crashes) and 30.87% of the severe injuries (wet pavement crashes). The nature of these crashes, especially for wet pavement, should be further investigated.

Table 5-43: Percent of Driver Involvements for Total and Wet Pavement Crashes by Skid Resistance

Skid Resistance	Percent of driver involvements			
	Total		Wet Pavement	
	Non-severe	Severe	Non-severe	Severe
1-25	34.35	28.65	37.72	30.87
26-34	0.61	0.49	0.93	0.61
35-44	57.07	61.05	55.16	60.53
>44	7.97	9.81	6.19	7.99
All numbers	100	100	100	100

5.4.1.2 Additional RCI Continuous Variable Transformations

The pavement condition (PAVECOND) field in RCI denotes a visual interpretation of the conditions of the road surface (FDOT, 2007). Where different lanes have different pavement conditions, the worst condition is recorded for the section. The categorization of the pavement condition rating was used in the group of the continuous values with good results. The joining of the poor and very poor categories was due to the sample size. Poor pavement conditions can contribute to loss of control and formation of water film on the roadway, leading to possible hydroplaning conditions. These conditions are important to test against the conditional probability of driver injury severity.

Table 5-44: Description of Pavement Condition Ratios (Source: FDOT, 2007)

Code	Description	Definition
1	Good	Good: First class ride with only slight surface deterioration.
2	Very good	Very Good: Only new or nearly new pavement.
3	Fair	Fair: Rutting, map cracking and extensive patching.
4	Poor or very poor	Very Poor: Virtually impassable. 75% or more deteriorated.
		Poor: Large potholes and deep cracks exist. Discomfort at slow speeds.

The sidewalk width and separation (SIDEWALK) is one characteristic that entails the pedestrian facilities for the location of the crash. It is also an indirect measure of the roadside clear zone, especially in urban areas. The values for the sidewalk width group were selected based in the standards in the Florida greenbook. The standard width of a sidewalk should be 5 ft when separated from the curb by a buffer strip, 4 ft may be considered when there are restrictions in the right of way. When sidewalks are adjacent to the curb, the minimum width is 6 ft. Unless there is high pedestrian traffic, wider sidewalks are not recommended to discourage higher bicycle speeds and possible increased traffic conflicts with vehicles at intersection and driveways.

A histogram analysis of the crash involvements proportions of non-severe injury involvements by sidewalk width group suggest a clustering of values 0, 5 and 6-8 ft. For the severe injury involvements the most common sidewalk widths were 3, 4 and 5 ft. Sidewalks less than 4 ft are considered non-compliant, older urban areas. Both 4 and 5 ft sidewalks are the minimums for urban areas with some buffer area between the road and the sidewalk. Sidewalks of 6 ft and more are those with no buffer or additional space needed due to high pedestrian volumes. Figure 5-8 below and Figure 5-9, page 190, clearly show a great difference between the portion of non-severe and severe injuries on locations with no sidewalks (<4 ft), which could be rural areas or suburban areas lacking pedestrian facilities.

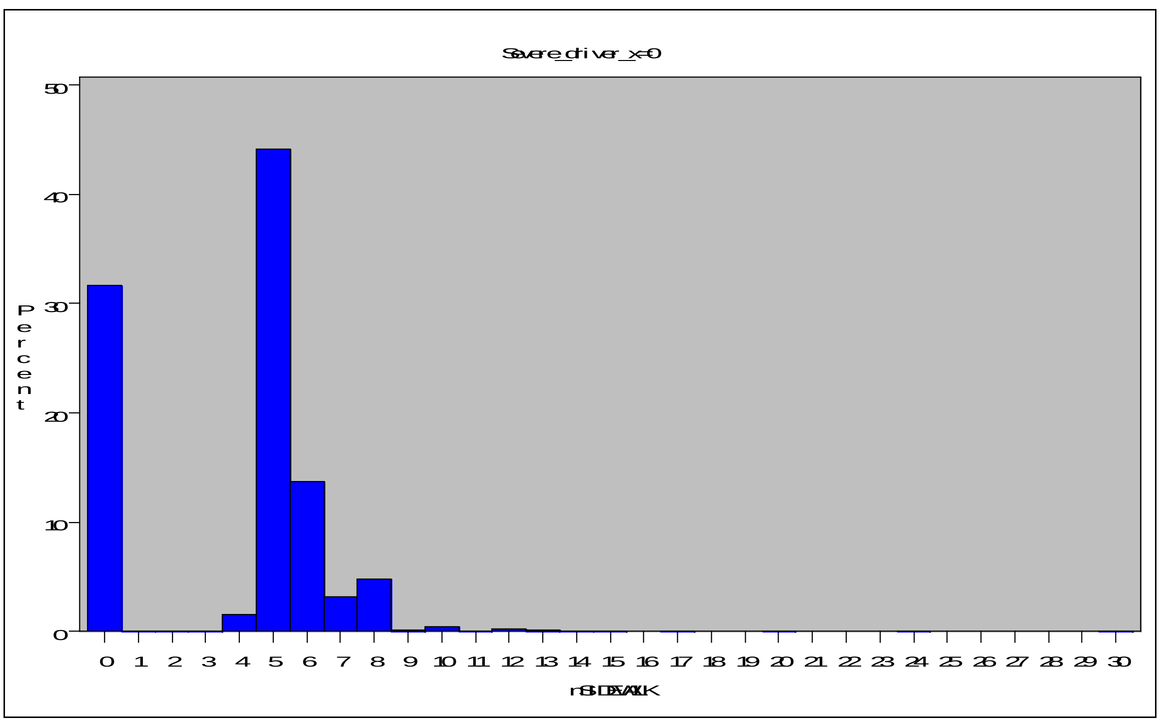


Figure 5-8: Distribution of Non-severe Involvements by Sidewalk Widths

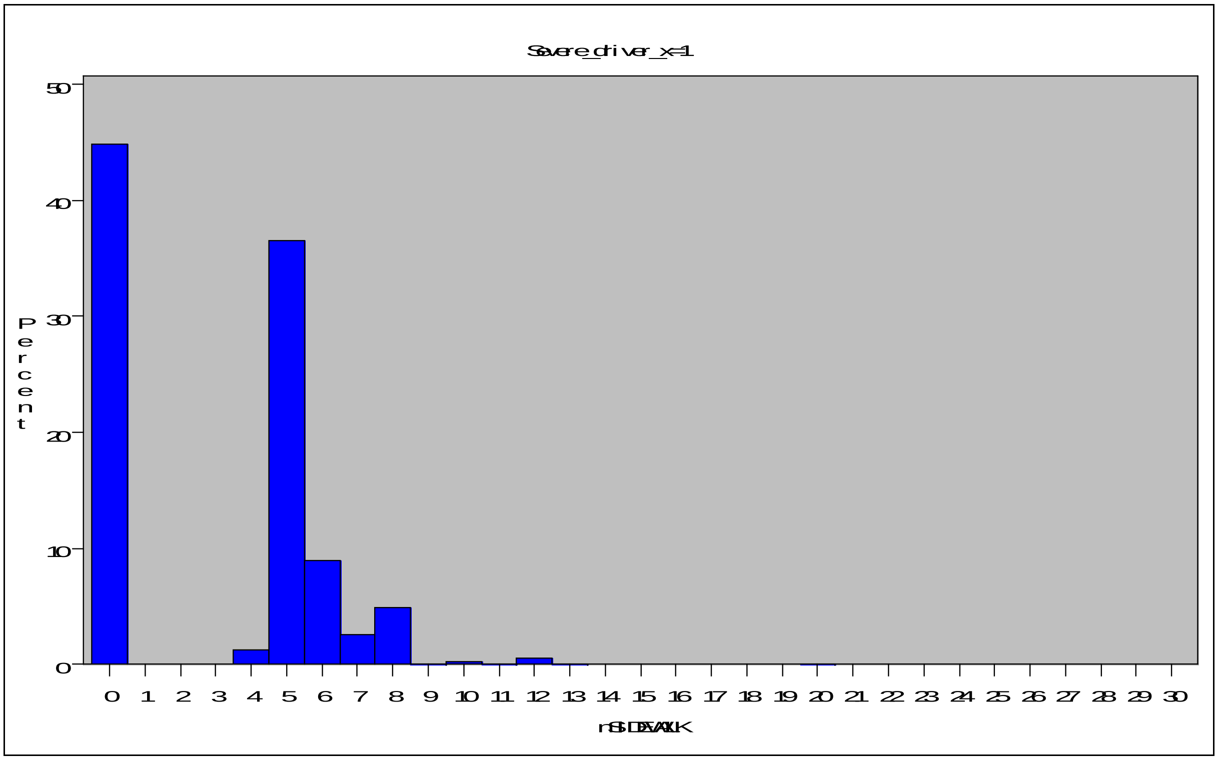


Figure 5-9: Distribution of Severe Involvements by Sidewalk Widths

5.4.2 Road Entity Models

The first six models of the final analysis performed well and had a distinctive information advantage over the previous models contained in Appendices E and F; Table 5-35 and Table 5-36. These models have remarkably better information performance, when comparing to the previous models. Most of the non-significant coefficients are for the *other* categories of the variables used in the model, which is acceptable. The following subsections present the model details by driver-, vehicle- and road-related variables. A complete list of variables descriptions for the final analysis is shown in Appendix C. The detailed models with coefficients and standard errors are presented in Appendix D.

Table 5-45: Goodness of fit for the Models Using the Stratified Dataset with Additional RCI Data

GOF Parameter	OVERALL	INTERS	SIGNAL	SEGMENT	PURE SEG	UNSIG
Number of Variables	38	24	20	32	22	16
Degrees of freedom	68	50	44	60	42	33
Marginally significant coefficients	5	3	1	10	1	3
Non-significant coefficients	7	5	7	7	10	3
Sample size	107449	55908	30832	69887	44818	25062
Response severe/non-severe ratio	6.09%	5.95%	5.23%	6.74%	6.68%	6.84%
AIC	41752.15	22153.74	11358.46	28556.32	17757.94	10792
Hosmer-Lemeshow p-value	0.7626	0.2168	0.6931	0.9695	0.5144	0.116
c value (area under ROC curve)	0.774	0.749	0.742	0.788	0.802	0.757
Percent Concordant	76.8	74.3	73.5	78.3	79.6	75.2
Adjusted R-squared	0.1883	0.1525	0.1314	0.213	0.2363	0.1749

The road entity models resulted in a rich model with driver, vehicle-collision and roadway-environment-related variables. The set of variables with the corresponding odds ratios are presented and discussed. The discussion is divided in these three major categories to allow for comparison of the six road entity models in an effective manner. Then, twenty models by crash type and road entities are discussed to complete the driver injury severity modeling presented in Section 3.6. Finally, additional discussion about the adequacy of the models is presented.

5.4.2.1 Driver-related Variables for Road Entity Models

The driver-related variables tested in the final analysis models are presented with their respective coefficient odds ratios in Table 5-46, page 192. Most of the important driver-related variables tested were found significant for any of the models. The odds ratios are used in this discussion as the best way to interpret the results of the logistic regression injury severity models, as discussed in Section 3.5.2. The values of the odds ratios represent a ratio of probabilities (holding all other factors constant) between a factor value and its base value. This ratio of probabilities refers to the chance of driver severe injury given a crash involvement occurs. The values presented in this section generally agreed with the empirical findings in past studies discussed in Chapter 2. Some of the driver-related variables have a high relative importance in the injury severity model, which will be discussed in a later section.

Table 5-46: Driver-related Variables Odds Ratios for the Final Analysis Models (Sample of Driver Involvements)

Variable	Level	Overall	Inters	Signal	Segment	Pure Segment	Non-signal
Driver Age Group (vs. 25-64 years)							
80-98 years	**5**	1.621	1.805	1.713	1.507	1.155†	1.922
65-79 years	**4**	1.422	1.542	1.426	1.410	1.253	1.665
20-24 years	**3**	0.779	0.755	0.695	0.800	0.803	0.823
15-19 years	**2**	0.767	0.793	0.789	0.752	0.750	0.799
Ejected (Yes/Partial vs. No)		4.270	4.209	4.281	4.381	4.533	4.170
Gender (Female vs. Male)		1.217	1.435	1.447	1.209	1.457	1.429
Seat Belt Used (vs. no)	**1**	0.303	0.344	0.348	0.302	0.322	0.346
Gender*Seat Belt Used	**1**	1.245			1.254		
Speeding (Unknown vs. not)	**2**	0.863†	0.610	0.724	0.825*	0.866†	0.607
Speeding (Yes vs. Not)	**1**	0.409	0.530	0.527	0.368	0.402	0.573
Contributing Cause (vs. no improper action)							
Other Contributing Cause	**4**	1.605	1.296		1.891	2.008	1.542
Aggressive Driving	**3**	1.748	1.432*		2.227	2.302	1.987
Alcohol/Drug involvement	**2**	1.593	1.622		1.786	1.645	1.858
At Fault driver (vs. not)		0.538	0.531	0.589	0.517	0.530	0.510
Red light running (vs. not)			1.333	1.439			
FL Resident (vs. not)		1.175	1.286	1.587			
Physical Defects (vs. not)		1.535	1.511	1.741	1.497	1.531	

Notes: * Effect is marginally significant (p<0.20); † Effect is not significant (p≥0.20)

Driver age has been proven to be a very important effect on the crash frequency and injury severity, as discussed in Chapter 2. The values of the odds ratios are consistent for all of the models presented here and similar to those on the exploratory analysis models. Very old drivers (ages 80 and over) have more than 150% probability of a severe crash when compared to middle-age (25-64 years) drivers in all the models. It is noteworthy that this coefficient (and odds ratio) was non-significant in the pure segment model and highest for the unsignalized intersection model. Very old drivers have more acute vision and gap judgment limitations which hamper their ability to cross intersections safely. This applies to both signalized and unsignalized intersections. Poor traffic signal coordination and severe unbalances towards major road traffic and limit gap availability of the unsignalized intersections.

Older drivers (65-79 years) follow the same trend as their older counterparts, but with slightly lesser odds ratios of severe injury, due to their more favorable physiological characteristics. The models agree with previous research that showed a correlation between intersections and older driver crashes. Younger drivers have odds ratios less than one, indicating a reduced probability (about 77% in the overall model) of severe injury when compared to middle age drivers, all else held constant.

Some of the studies discussed in Chapter 2 focused on the different injury severity experience of males and female drivers. The differences are mainly due to increased physiological resistance of male drivers to injury. Female drivers are 121% more likely to sustain severe injury as a result of crashes on high-speed multilane roads. This negative effect is more acute at intersections, with a 143% chance of severe injury when compared to males, all else held constant. The odds ratios do not vary consistently among the models, for example the segment and overall models have lower odds ratios than any of the other road entities that make up these combined models. The interaction effect of gender and safety equipment (seat belt use) shows the effect of females using seat belts compared to males not using seat belts. Even with this disadvantage, females showed a 125% chance of severe injuries compared to male drivers. This interaction was only present in the overall and segment models and explains the lower odds ratio for the gender main effects variable discussed previously. However, driver ejection did not present an interaction with gender or age. The ejected variable odds ratio shows that drivers are ejected (partially or totally) from their vehicle are 427% more likely to sustain severe injury than those who are not. These models show a better picture of the magnitude and related factors to female driver injury severity.

Behavioral differences by gender may have some effect on injury severity; however, those could not be directly correlated (as interactions) in the models presented here. Seat belt usage is the single most important positive effect on injury severity, as indicated by its odds ratio. Drivers using seat belts are only 30% likely to sustain severe injury when compared to those who do not, holding all other factors constant. Driver ejection, although a post-crash event, is important for these models to complement the seat belt use effect. The issue of seat belt use over reporting was discussed earlier in this chapter and may affect the validity of the magnitude of the positive effect of seat belt usage. Driver behavior is the most important contributor to their injury severity outcomes in crashes on high-speed multilane roads. Engineering countermeasures should focus on improving the decision capabilities and focus on the road of drivers in these busy corridors.

Drivers at fault were found to have a 53% likelihood to sustain severe injury when compared to those not at fault. This relation was almost constant, including the signalized intersections, where the red light running drivers are 143% more likely to have severe injuries than non red light running drivers. Red light running crashes are generally the most severe crashes for signalized intersections and have become a very important target of engineering and law enforcement countermeasures in Florida. Speeding drivers are 40% as likely as others to sustain sever injury. This variable has a degree of indeterminacy due to the unknown values for 24% of the involvements in the stratified sample and the estimation of speed by the police officer. However, due to its importance, it was tested and the unknown level was non-significant in the overall model, suggesting that this trend will be true if all estimated speeds were known. The estimation effect is less likely to affect the speeding variable due to the training and experience of officers in judging speeding driver behavior. The contributing cause factor showed the expected higher chance of aggressive drivers (159%) to sustain severe injury as a result of crashes in high-speed multilane roads. This effect is more pronounced on unsignalized intersections, were poor gap availability may lead to aggressive driving compounded by aggressive drivers on the major road that do not keep a safe distance for reaction to unforeseen perpendicular crossing events.

Some relatively less important driver variables were not dropped from the model when additional road variables were introduced. This improved the information available on the models. Drivers who are Florida residents are 117% as likely as the non-residents to sustain severe injury in crashes on high-speed multilane arterials, holding all other factors constant. This is an important result out of concern for the effects of the millions of visitors that drive in Florida each year on the safety performance of the high-speed multilane corridors. Only 4.61% of the total drivers involvements analyzed were classified as non-resident. It was observed that non-resident drivers were involved in 4.34% of the severe injuries (see Table 5-47, page 196). Only 1.34% of the severe injuries involved a non-resident driver at-fault. No major difference was evident between non-resident at-fault drivers (44.5%) and the overall at-fault rate in severe injuries (40.6%, see Table 5-9). Past research by Abdel-Aty et al. (1999) showed that non-resident driver crash involvements by age followed a similar trend than for the Florida resident drivers. The results of the injury severity models demonstrate that non-resident drivers are not a major safety concern for high-speed multilane arterials in Florida. Drivers suffering from physical conditions (defects) such as fatigue and eyesight problems are 153% as likely to sustain severe injury as healthy drivers, holding all other factors constant. The signalized intersections represent the biggest challenge to these drivers, were the odds ratio for the physical defect variable increased to 1.74, suggesting that drivers in bad physical condition fare worse in crashes at or near intersections.

Table 5-47: Driver Residence and At Fault in Severe Crash Involvements

At Fault Driver	Residence_Code		Total	Percent
Frequency (Percent)	Non-Florida	Florida		
No	200 (3.00)	4542 (68.22)	4742	71.22%
Yes	89 (1.34)	1827 (27.44)	1916	28.78%
Total	289	6369	6658	100.00%
Percent	4.34%	95.66%		

5.4.2.2 Vehicle- and Collision-related Variables for Road Entity Models

The vehicle- and collision-related variables are important in contributing to the injury severity characteristics found at the high-speed multilane roads. The variables found significant in all models are harmful event (collision type) and type of vehicle. The interpretations of these variables complemented the vehicle maneuver and private vehicle use variables in some models.

Important types of crashes, such as off road, work area-related, and multivehicle, were significant in the overall model; the latter proves to be most appropriate in terms of coefficient significance (see Table 5-48). The odds ratio values agreed with previous empirical results.

Table 5-48: Vehicle- and Collision-related Variables Odds Ratios for the Final Analysis Models

Variable	Level	Overall	Inters	Signal	Segment	Pure Segment	Non-signal
Harmful Event Group (vs. Rear-End)							
Other	**7**	1.097	1.091*	0.982†	1.132	1.149	1.101†
Fixed Object	**6**	1.810	2.174	2.578	1.840	1.879	1.636
Sideswipe	**5**	0.779	0.724	0.540	0.836*	0.904†	0.810†
Left Turn	**4**	2.242	2.228	2.304	2.026	2.306	1.806
Angle	**3**	1.784	1.812	1.649	1.790	1.818	1.754
Head-On	**2**	2.875	2.263	2.188	3.211	4.054	2.068
Vehicle Maneuver (vs. Straight Ahead)							
Other	**4**			1.100†			
Left Turn	**3**			1.258			
Slowing / Stopping	**2**			0.760			
Vehicle Type (vs. Automobile)							
Other	**5**	0.756	0.654	0.597*	0.798*	0.824†	0.708†
Bicycle and motorcycle	**4**	1.050†	1.273	1.186†	0.965†	0.926†	1.352*
Trucks and buses	**3**	0.357	0.325	0.366	0.362	0.371	0.273
Van, Light Truck, Pick up	**2**	0.820	0.780	0.811	0.830	0.872	0.742
Point impact (driver side vs. not)		1.091†	1.187*	1.524	1.094†		1.595
Point impact*Speeding	**2**	1.240*	1.457		1.093†		
Point impact*Speeding	**1**	1.412	1.507		.		
Off Roadway (vs. not)		0.613			0.739	0.716	
Off Roadway*Speeding	**2**	0.783*			0.896†	0.944†	
Off Roadway*Speeding	**1**	1.289			1.517	1.388	
Off Roadway*Multivehicle	**1**	2.033					
Work Area (Entered vs. none)	**3**	0.826*			0.829*		
Work Area (Nearby vs. none)	**2**	0.750			0.742		
Private vehicle use (vs. not)					1.565		
Private veh*Avg Truck Factor					0.975		
Multivehicle (vs. single vehicle)	**1**	0.469					
Intersection*Multivehicle	**1**	1.476					

Notes: * Effect is marginally significant (p<0.20); † Effect is not significant (p≥0.20)

The odds ratios just shown in Table 5-48, page 197, indicate that only sideswipe collisions are less likely (22.1% less) to result in driver severe injury when compared to rear-end crashes. The odds ratios for collision types significantly varied among the road entity models, confirming the reasoning for the proposed modeling scheme by crash types presented later in this chapter. Drivers' involved in head-on crashes were 287% as likely as those involved in rear-end crashes to sustain severe injury, the highest odds ratio among collision types. Head-on crash involvements are most severe in pure segments (cross-over of the median) with a 405% chance of severe injury when compared to rear-end involvements. Even when the operating speeds are lower than in limited access facilities, the severity of this type of involvements should be taken into account when designing and improving these arterial corridors. The left turn crashes are of great concern because they are more frequent (6% of the total) than head-on crashes on high-speed multilane roads (see Figure 5-10). Drivers involved in left turn crashes are 224% as likely to sustain severe injury as those in rear-end crashes. On the other hand, drivers involved in angle crashes are 178% as likely to sustain severe injury as those involved in rear-end crashes. This figure is alarming because angle crashes are more than double as frequent as the left turn crashes.

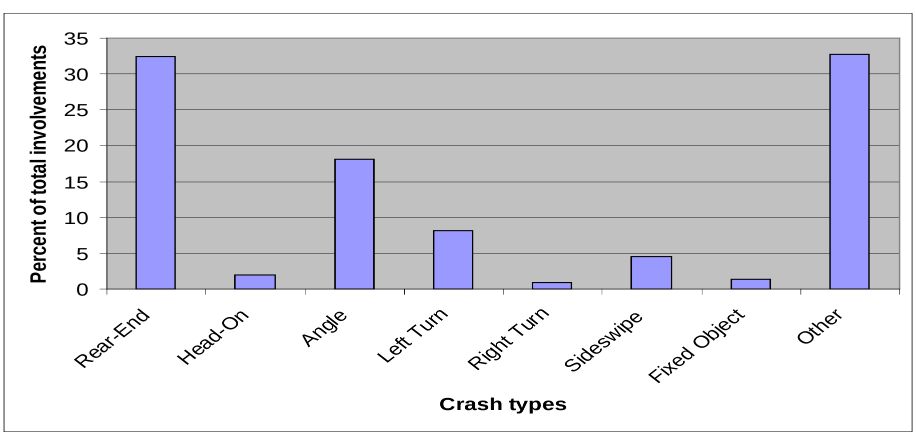


Figure 5-10: Collision Types on High-speed Multilane Roads

Another group of crashes that are important in high-speed multilane arterials are those involving fixed objects. Drivers involved in these types of collision are 181% as likely to sustain severe injury as those in rear-end collisions. Although this group is not a major component of the total number of crashes (see Figure 5-10), it is directly affected by the roadside design characteristics. The lack of uniformity in access management and limitations of right of way in urban areas constitute major design challenge for new and reconstructed arterial corridors. The level of disaggregate analysis in the injury severity models allows to pinpoint major factors that affect the severity performance in high-speed multilane roads. The models by crash types will uncover additional contrasts discussed in the latter part of this chapter.

The vehicle maneuver variable for left turn movements (vs. straight ahead) has a chance of 125% to sustain severe injury compared with the movement correlated with rear-end crashes. The differences between the left turn (harmful event) and left turn vehicle maneuver collision odds ratios suggests that in general drivers making a left turn are less likely to sustain severe injury than those hitting the left turning vehicle in signalized intersections. The left turn phasing in signalized intersections may affect the severity outcomes for the left turn crashes. It seems that he permissive left turn phasing is an underlying factor, since this variable was only significant for the signalized intersection model. Previous research has shown that the left turn phasing in the major road is a contributing factor to reduced crash frequency along arterial corridors (Abdel-Aty and Wang, 2006). In that particular study, the negative binomial crash frequency model accounted for the signalized intersection spatial correlation along the corridor. In another study, Wang and Abdel-Aty (2006) found that for rear-end crashes would increase when both left turn phases in the major road are protected in crash frequency models accounting for spatial correlation of signalized intersections in arterial corridors. The safety offset of reducing the most severe crash involvements is an increase in certain crash types that are less severe.

The type of vehicle is an important factor affecting the injury severity of drivers involved in crashes. The changing nature of the vehicular fleet and roadway design, especially roadside appurtenances such as guardrails and crash cushions, may affect the safety performance of the high-speed multilane corridors. Drivers of vans, light trucks (LTV's) and pickups involved in crashes at high-speed multilane roads are 82% as likely as passenger car drivers to sustain severe injuries. The adverse effects of the LTV's visibility obstruction on following passenger cars in rear-end and angle crashes resulting in fatal injuries and the differential impact energy effect on fatal angle collisions have been studied by Abdel-Aty and Abdelwahab (2003 and 2004).

Drivers of heavy trucks are 35.7% as likely as passenger car drivers to sustain severe injury from a crash involvement. This heavy truck driver effect is more acute in unsignalized intersections (27.3% as likely as passenger car drivers to sustain severe injury). Duncan et al. (1998) found that impact speed differential and passenger car rear impact were contributing factors to increasing incapacitating injury and fatal injuries in car-truck two vehicle crashes on divided highways in North Carolina. Additional research is needed to determine the specific factors affecting car-truck crashes at unsignalized intersections. On the other hand, private vehicle use (almost certainly not heavy trucks) was found to increase the likelihood (compared to non-private vehicles) of severe injuries in segments, while the interaction with the average truck volume factor showed decreased likelihood with increased truck volumes. These additional factors and their possible relation to type of vehicle, land use and travel choice needs to be further investigated.

The bicyclists and motorcyclists' injury severity coefficients are significant only in the intersection model. This point to a need of increased efforts to reduce traffic conflicts at intersections, reduce red light running and other situations were these vulnerable road users are at a disadvantage with the other vehicle types. Research on motorcycle and bicycle crash rates has been hampered by a lack of exposure data, prompting new data integration strategies in the Florida 2006 SHSP (FDOT, 2007). The increase in motorcycle fatal crashes as a proportion of all motor vehicle crashes in Florida and the high proportion of intersection crashes indicate the potential for safety problems on intersections in relation to these vulnerable drivers.

The point of impact variable is significant only for the signalized intersections, which points out that drivers of vehicles hit on the driver's side are 152.4% as likely as those of vehicles with impacts on other locations, all else held constant. These results contrast with the results of the injury severity analysis using data from all non-limited access roads in three counties in Central Florida (Abdel-Aty, 2003). In that study, the point of impact variable (driver side vs. non-driver side) was significant in both the road segments and signalized intersections models. The higher likelihood of severe injury for drivers of vehicles hit on their side agrees with the previous study. The interaction of point of impact and speeding helps explain this difference. Drivers speeding at or near intersections and hit by their side of the vehicle are 150.7% as likely as the other drivers of sustaining severe injury, all else held constant. If we compare the speeding coefficients found in that previous study with the speeding coefficient found in the models of the final analysis, it seems that the likelihood of increased injury severity is higher for high-speed multilane corridors. It can be theorized that driver injury severity in high-speed multilane roads are more affected by speeding behaviors and intersection locations than other types of roads (i.e.

two-lane roads). Additional research could test the validity of this claim and describe its implications.

Off roadway crashes (part of lane departure crashes) have been identified by the SHSP as a serious safety hazard (FDOT, 2007). The majority of the lane departure crashes occur on limited access roads and two-lane roads. However, the significance of the fixed objects and off roadway coefficients indicates that these affect the safety performance (in regards to driver injury severity) of the high-speed multilane roads. The drivers in off roadway involvements are 61.3% as likely to receive severe injury as the other drivers, all else held constant. This seems as good news in terms of high-speed multilane roads. Two interaction coefficients, of which one is only significant in the overall model, present a more complete picture of the off roadway crash involvements. Drivers speeding and involved in off roadway crashes are 128.9% as likely as other drivers to sustain severe injury, all else held constant. In addition, drivers involved in multivehicle off-roadway crashes are 203.3% as likely as other drivers to sustain severe injury, all else held constant. These off-road crash interactions emphasize the role of speeding as a safety hazard in high-speed multilane roads and points to a possible relation between driveway crashes (multivehicle and off road) and increased injury severity. The implications of these results are further discussed in Chapter 6.

The work area variable significance in the overall models was not expected. This denotes the importance of exploring the safety effects of work zones in high-speed multilane roads. The effect on driver injury severity is positive (OR=0.75) for the crashes entering the work zone when compared to crashes outside work zones. Meanwhile, crashes inside the work area did not have a significant effect on driver injury severity. This result suggests that crashes inside work areas could be more severe than those in the transition area (taper). In the rural and urban areas models, there are major differences with these results. A possible avenue of research is to investigate the relationships between speed limits, land use and work area crash frequency and severity.

The single vehicle crashes tend to produce more severe injuries in high-speed multilane arterials. Drivers involved in multivehicle crashes are 46.9% as likely to sustain severe injury as the drivers involved in single vehicle crashes. Single vehicle crashes constitute less than 9% of the total crashes and are still shown as a significant factor contributing to injury severity. Multi-vehicle involvements at (or near) intersections are 147% as likely to result in severe injury as the single vehicle involvements at or related to intersections. The multiple interactions of factors with intersections underscore the importance of these road entities in the analysis of the multilane arterial safety performance.

5.4.2.3 Roadway- and Environment-related Variables for Road Entity Models

Many of the variables discussed previously have indirectly affected the roadway design in high-speed multilane roads. However, the roadway variables are the most important for a study seeking to find contributing factors that describe the safety performance (in terms of driver injury severity) of multilane road design characteristics. These are expected to allow researchers find ways to improve the design, maintenance, operation and safety performance of the arterial corridors that are under increased pressure to sustain the mobility needs of a growing population. The group variables found significant in at least one of the road entities are shown in Table 5-49, page 204. A large number of roadway-related variables was found significant even with the presence of a complete set of driver vehicle- and crash-related variables. This is one of the main achievements of a painstaking data preparation and model development effort.

Table 5-49: Roadway- and Environment-related Variables Odds Ratios for the Final Analysis Models

Variable	Level	Overall	Inters	Signal	Segment	Pure Segment	Non-signal
Speed limit (40-45 vs. other)	**1**	0.676	0.657	0.709	0.657	0.677	0.622
***adt* per Lane (thousands)**		0.972			0.960	0.948	0.972
Avg Truck Factor (percent)		1.011			1.034	1.012	
High Mast (full vs. none)	**Y**	1.331			1.403		
High Mast (partial vs. none)	**P**	3.506			3.710		
Traffic Signal (vs. other)	**3**	1.128			1.236		
Stop/Flashing (vs. other)	**2**	0.997†			1.319*		
Non applicable (vs. class 2,3,4)	**9**	1.030†	1.009†	1.040†			
Access class 7 (vs. class 2,3,4)	**7**	0.781	0.677	0.622			
Access class 6 (vs. class 2,3,4)	**6**	0.833	0.756	0.649			
Access class 5 (vs. class 2,3,4)	**5**	0.879	0.847	0.836			
Urban area (vs. Rural)		0.879	0.744	0.808	0.816	0.916†	0.649
Curb Shoulder (vs. Paved)	**3**	1.089	1.114	1.162			
Unpaved Shoulder (vs. Paved)	**2**	0.967†	0.973†	0.954†			
Lane width (vs. 11 ft ≤ width ≤ 12 ft)							
Lane width <10 ft	**4**	0.827	0.766	0.812	0.843		
10 ft ≤ lane width < 11 ft	**3**	0.815	0.802	0.677	0.874*		
Lane width > 12 ft	**2**	0.810	0.850	0.801	0.818		
Roadway Curve (vs. non curve)		1.306			1.391	1.339	1.510
Sidewalk width ≥ 6 ft (vs. <4 ft)	**3**	0.791	0.723	0.665	0.801	0.831	0.762
4 ft ≤ Sidewalk < 6 ft (vs. <4 ft)	**2**	0.851	0.774	0.696	0.890	0.910*	0.833
Full Non-High Mast (vs. none)	**Y**	1.129	1.162		1.121*	1.053†	
Partial Non-High Mast (vs.none)	**P**	0.821	0.873		0.837	0.770	
Type Friction Course (vs. FC-2)							
FC-3, FC-6 or not applicable	**9**	0.975†	0.995†	0.916†	0.992†	0.948†	
Friction Course type 5 (FC-5)	**5**	0.831	0.726	0.539	0.903†	0.935†	
Friction Course type 4 (FC-4)	**4**	0.918	0.976†	0.988†	0.929*	0.888	
Friction Course type 1 (FC-1)	**1**	0.736	0.737	0.687	0.754	0.723	
Intersection		0.831*					
Intersection*Urban area		0.862					
Skid Res. (1≤FN<35 vs. FN≥35)		1.198	1.143		1.256	1.305	
Urban area*Skid Resistance		0.919*	0.969†		0.868*	0.825	
3 or more Aux Lanes (vs. none)	**3**				1.075†		1.190*
2 Auxiliary Lanes (vs. none)	**2**				0.882		0.860*
1 Auxiliary Lanes (vs. none)	**1**				1.054*		1.136
Weekend (vs. Weekday)		0.906					

Notes: * Effect is marginally significant (p<0.20); † Effect is not significant (p≥0.20)

The key advantages of the overall model to analyze the driver injury severity resulting from crash involvements on high-speed multilane arterial corridors are realized in the roadway-related variables. The implications of these results are discussed further in Chapter 6. One of the most important road design parameters is the speed. Speed limits generally reflect a decrease of 5 to 10 mph below the speed limit. In regards to driver injury severity, only speed limits 40 to 45 miles per hour resulted in significant coefficients. Drivers involved in crashes on multilane roads with speed limits 40 to 45 miles per hour are 67.6% as likely to sustain severe injury as drivers in non-limited access roads with other speed limits. This applies to the drivers on minor roads (lower speed limits) as well as those on higher speed roads, up to 70 mph. Recall that a 70 mph speed limit could be recorded when the crash is related to a ramp intersection with a limited access road such as the Florida Turnpike. These represent a very small minority of the cases. Speed differentials are correlated to increased crash severity (Wilmot and Khanal, 1999); thus the reduced severity on the most common speed limits is expected when compared to the roads with smaller or higher speed limits. It also points to a possible increase of severity on multilane roads with speed limits of 50 mph and over. This should be the subject of further severe crash frequency research. The larger odds ratio for the signalized intersection suggests that signalized intersections are more sensitive to the major road speed limit than other road entities.

The traffic volume per lane is one of the most important contributing factors to driver injury severity. For each unit increase of *adt* per lane, there is a 97% conditional probability severe injury compared to the previous unit of *adt* per lane and holding all other factors constant. Each unit of *adt* per lane consists of 1,000 vehicles, as explained earlier in this chapter. As traffic volume increases, severe injuries tend to decrease. This relationship between congestion and safety has been proven in previous research (Shefer and Rietveld, 1997). Another important result is that the joined intersection (combined signalized and unsignalized) did not improve with a negligible decrease of 2 in the AIC value after forcing the *adt* per lane variable into the model. Since *adt* per lane was significant in the unsignalized intersection model, it was also tested in the signalized intersection model. The signalized intersection model worsen (AIC increased by 10) when the variable *adt* per lane was forced into the model after the variable did not enter the model by the stepwise method. The coefficient of the variable was highly insignificant (p-value=0.4064). Recent research by Wang et al. (2006) and Abdel-Aty and Wang (2006) showed that the natural log of the *adt* per Lane was found significant in the crash frequency models for signalized intersections in arterial corridors. The change in functional form of the *adt* per lane variable was tested with similar results. Additional tests and discussion about this situation are presented at the end of this chapter.

The average truck factor is the proportion of heavy truck volume on roads. As the truck traffic increases, there is an increase, albeit smaller than for other factors, in the severe injuries on high-speed multilane roads. There is a tendency of higher severe injuries in heavy business and industrial areas. Thus, this effect is expected to be correlated to the land use. Special consideration should be given to the provisions for deceleration, acceleration, and storage lanes in commercial or industrial areas with significant truck/bus traffic.

There are two variables regarding roadway lighting: LIGHTING refers to the density of high mast luminaries per mile, LIGHCDE refers to the density of non-high mast luminaries per mile. There is a strong tendency in Florida to place high mast lighting mainly in areas that require a large spread of illumination, such as interchanges and limited access facilities exits. The rest of the locations use conventional light poles, especially in urban areas. The results of the models indicate that the high mast is significant only in the combined segment models, which includes unsignalized intersections. Many isolated rural areas with interchanges are more likely to have high mast lighting, but not enough to have a significant proportion of all of the unsignalized intersections. In the case of the non-high mast lighting, the variables were only significant in the intersections, segment and pure segment. The non-significant results in the signalized intersections model are suspected to be related to the differences in land use. Unsignalized intersections were expected to have non-significant effects because of the tendency to have lower lighting levels in the rural areas. When comparing the odds ratios of the high mast and non-high mast lighting, there is an increased benefit for the partial non-high mast lighting (odds ratio 0.821 for overall). The negative effects (odds ratios greater than one) of the non-high mast lighting has to be interpreted taking into account the types of locations were these luminaries are located. Rural locations have a clear tendency to higher injury severity. However, the very high odds ratios (3.506) for the partial lighting may require special attention as a possible group of hazardous locations in high-speed multilane roads. There is a clear advantage to locations with partial non-high mast lighting. The high mast lighting negative effects are likely caused by the rural intersections locations were high masts tends to be placed. However, there is a large difference between the partial and full lighting density, which suggest the large benefits of additional lighting in these areas. Again the overall model proved adequate with significant contributing factors that are consistent with the values in the individual models in which these were also found significant.

Access Management classes in a multilane arterial corridor are a measure of the type of median and proximity of median openings, access points and intersections. Classes 2 to 4 are characterized by a signalized intersections separated by ½ mile or more, while for classes 5 to 7 the separation is reduced to ¼ mile for roads with speed limits 50 mph or greater. Compared to classes 2 to 4 with longer separation distances, classes 5 to 7 exhibited odds ratios less than one. Class type 7 seems to exhibit the most benefit (odds ratio=0.781), followed by class 6 (odds ratio=0.833) and class 5 (odds ratio=0.879). It is important to reiterate that the logistic regression represents a conditional probability given a crash occurs, which does not necessarily traduces into increased severe crash frequency. The lower speed limits and higher traffic congestion are expected to lower the risk of severe crashes in these areas. The type of median opening becomes an important effect on the injury severity when the signal spacing is reduced to ¼ of a mile. The distance measurement of median (unsignalized intersection) separation and intersection spacing are taken from the center of the intersections. The negative effects of closely spaced intersections has been suggested by previous research, but the effect of median opening types has seen limited study in part due to the difficulties raised by the experimental design of before after-studies of major corridor improvements. The directional median openings restrict crossing maneuvers that increase the likelihood of angle crashes, which tend to be severe.

When the overall model was built using urban area crashes only, the odds ratio of class 5 remained practically the same, while the odds ratio of class 7 raised by 14.61%. In urban areas only, the odds ratio of classes 5 to 7 are 0.873, 0.795 and 0.667, respectively. This demonstrates the benefits of the median access restrictions and longer spacing between access points and intersections. Since each class has two access point separations options by speed limit it is important to determine whether other factors affect the safety performance of road sections with closely spaced intersections. Further research is needed in this area.

The traffic control variable shows how severe injuries outcomes may be affected by the type of traffic control of intersections. There is a trend of lower severe injury risk at signalized intersections, but this coefficient was not significant mainly due to the issues discussed previously with the signalized intersections in rural and urban areas. However, it is clear that the stop controlled intersections represent a higher risk (odds ratio=1.128) when compared to other intersections and road segments.

The land use is the single most important factor that has indirectly affected these models. The close agreement of the access class variable odds ratios and the rural urban variable in Table 5-49, page 204, are yet another proof of that phenomenon. Access classes 5 to 7 are expected in urban areas. Clearly, the urban area exhibits lower conditional risk of severe injury, holding all else constant (odds ratio=0.879). The non-significance of the coefficient in the pure segment model suggests that the major differences between rural and urban areas in terms of injury severity are in regards to their intersections and access points.

The type of shoulder variable showed the effect of this important component of the road in the context to its correlation with land use. The base variable, paved shoulder exists in both rural and urban (or suburban) land uses. The unpaved shoulders are usually found in rural areas, but it seems that not with enough frequencies to make it a significant factor in the driver injury severity. However, the curb shoulder type suggests higher severity in crashes (odds ratio=1.089) than the paved shoulder. Sections with curbs usually have narrow shoulders and curbs themselves may become a barrier at high speeds. This is likely to affect the risk of severe injury in crashes, given a crash occurs. This effect was significant for the signalized intersections which lead to possible relationships between the degree of development, traffic conflicts and severity of crashes, especially at urban signalized intersections.

The lane width coefficients were compared to a base value of lanes between 11 and 12 ft wide, which are the most common used lane widths. The odds ratios of lanes less than 10 ft (0.810) and between 10 and less than 11 ft wide (0.815) are practically equal. Meanwhile, for lanes more than 12 ft width (odds ratio=0.827) there is a perceived benefit of a wider lane on multilane arterials. The positive effects of increased lane widths on crash frequencies on an arterial corridor have been discussed previously in a study by Abdel-Aty and Radwan (2000). Another study by Karlaftis and Golias (2002) pointed out the positive effects of lane widths on crash rates of multilane roads.

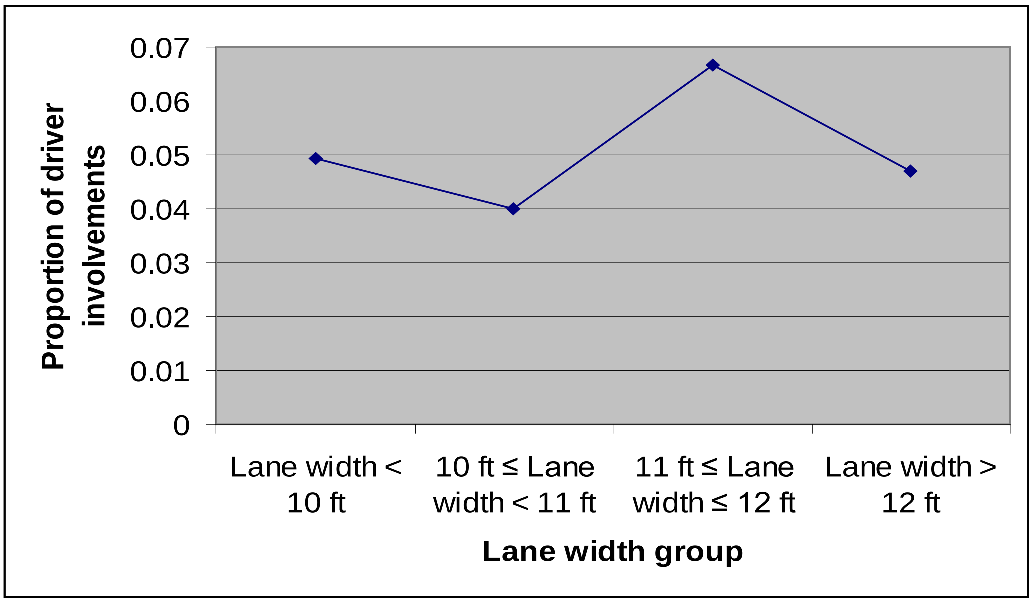


Figure 5-11: Lane Width and Severe Injury Proportions on High-speed Multilane Roads

Preliminary analysis of the severe involvements suggested a positive effect of lane widths larger than 12 ft, or between 10 and less than 11 ft (see Figure 5-11). This might be explained by the urban settings where lower lane widths are used with lower operating speeds. The lanes less than 11 ft wide are generally found in urban areas. In an overall model using urban area crashes only, there is an increased benefit of lanes wider than 12 ft (odds ratio=0.806 vs. 0.855 and 0.866 for the other two lane width groups). In the overall model in the rural area only, there is an increased benefit for lanes 10 or less than 11 ft wide (odds ratio=0.675) compared to the lanes more than 12 ft wide (OR=0.882). For lanes less than 10 ft wide, the results agree with past studies (basis of the Florida greenbook) that suggested a lower bound of 10 ft. On the other hand, the odds ratios for lane widths between 11 and 12 ft, suggests that these widths may not be enough for the high-speed multilane arterials in urban areas. Probable causes may include a changing vehicle fleet (SUV's), higher truck volumes and high operating speeds.

Roadway curves have a negative effect on driver injury severity (OR=1.306). However, on high-speed multilane corridors this effect also varies in rural and urban areas. In the urban area model there is a much greater negative effect (OR=1.531) vs. in the rural area (OR=1.197). This result can be understood if the conditions on roadway curves in urban arterials are analyzed. Increased access point conflicts, visual distraction and difficulties in negotiating consecutive lanes changes from a driveway to a turn lane are conditions likely to increase the severity of crashes, making angle crashes more frequent in urban curves.

Sidewalks have proven safety benefits for pedestrians by providing a secure path apart from vehicular traffic. Every model showed significant effects of the sidewalk width, which make this a strong contributing factor only emulated in achieving significance in every model by land use and speed limit. Its main benefit is seen in the unsignalized intersection model because the difference between the 6 ft and over width and the 4-6 ft width group is greater (OR 0.7662 vs. 0.833). Having wider sidewalks increases the chance of improved roadside clear zones and less visual obstructions. The odds ratio of sidewalks with minimum length (0.851) shows benefit in comparison with no sidewalk, but sidewalk widths of 6 ft and over showed increased benefit (OR=0.791). There is a clear benefit to wider sidewalks than the minimum standard, having substantive safety rather than nominal safety.

The presence of a friction course serves to reduce severe crashes in wet pavements by increasing skid resistance. The relationships between these two variables were discussed in Section 5.4.1.1. There is a tendency to increased sever injuries in sections of road with older friction courses. Newer open grade friction courses (such as FC-5) was found to have a positive effect (OR=0.831) when compared to the older types (FC-2). There is an increased benefit in an older friction course (FC-1, OR=0.736) which needs to be further investigated. The urban and rural area models points to some fundamental differences in the effects of this variable.

The intersection variable had a marginally significant positive effect, probably due to the urban and rural area differences between the road entities. The interaction with the land use clearly shows a positive effect of intersections in urban areas (OR=0.862). In rural areas, there is no significant benefit, while in urban areas, there is a significant positive effect (OR=0.621). Further research is required to investigate these differences.

Increased skid resistance numbers provide better friction to perform stopping (or swerving) maneuvers in wet pavements. The negative effect on driver injury severity (OR=1.198) comes unexpected. The interaction effect with land use indicates that there is a positive effect in urban areas (OR=0.919) although marginally significant. Both coefficients compared skid resistance numbers equal or greater to 35 (acceptable by FDOT standards) to those lower than 35. Analysis of an overall model using wet pavement crashes only was warranted since the benefits of added friction are expected on wet pavements. The model resulted in the same sign for these coefficients. Wet pavement crashes were analyzed by skid resistance numbers (see Figure 5-12, page 213). There is a clear trend of lower severe involvements ratio to total involvements for higher skid resistance numbers (>44). However, in the rural areas, there is a significant increase for higher skid resistance numbers. The interaction lines before 35 are parallel, denoting a correlation between urban and rural areas for wet pavement crashes. This correlation is not as clear in the distribution of all crashes (see Figure 5-12). There is some evidence of a systematic positive benefit in urban areas and negative effect in rural areas. However, in the overall models in rural and urban areas, the odds ratio remained greater than one.

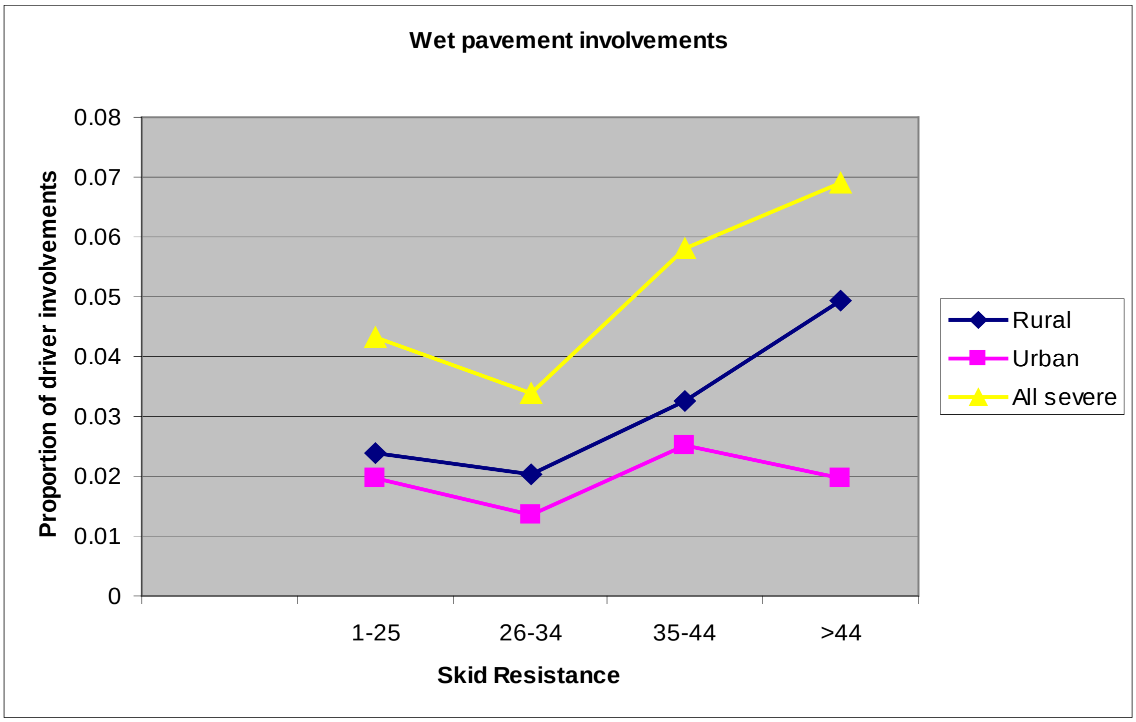


Figure 5-12: Skid Resistance and Severe Injury Proportions for Wet Pavement Crashes on High-speed Multilane Roads

According to the FDOT Skid Hazard Manual, for a skid overlay to be effective in reducing crashes, a significant portion of the crashes should be occurring on wet pavement. Evaluation of past skid hazard improvement projects, found that when at least 25 percent of crashes occurred during wet weather, crashes were more likely to decrease after the improvement. Additional research should compare road sections with higher proportions of wet pavement crashes with the rest to determine the relationships (if any) between these sites and other road characteristics that may increase (or decrease) the risk of severe crashes. Moreover, this variable could be picking up some driver effects, similar to the overconfidence of the offset hypothesis explained earlier. In a recent study Mannering (2007) found that drivers who believe pavement quality on Indiana interstates is good or very good were more likely to drive at higher speeds. The effect was higher on speed limits 55 mph, compared to speed limits 65 and 70 mph.

The auxiliary lane variable was only marginally significant in combined segment and unsignalized models. There is a tendency of decreased injury severity when two auxiliary lanes are present. This may be correlated with the presence of right and left turn lanes in unsignalized intersections. There is no evidence of a strong effect of the presence of auxiliary lanes in multilane arterials. This might be due to the weight of other variables such as Access Class, which are correlated to the auxiliary lanes on medians.

The day of week variable compared weekend vs. weekdays. The overall model reflected a positive effect (OR=0.906) for drivers on weekends compared to the weekdays. Local experience suggests increased traffic volumes during weekends on multilane arterials during most of the day. A more uniform traffic volume across time of day reduces operating speed, decreasing the chance of severe crashes.

5.4.3 Crash Type Models

The driver injury severity analysis was developed using five crash type models. The combination intersection model was not deemed necessary in the original plan and the final model analysis confirmed this assertion. The same variables available for the road entity models in order were tested to make both sets of models fully comparable. The main crash types reported on high-speed multilane roads were included in this analysis. Rear-end, angle, and left turn crashes are the most common crash types on high-speed multilane roads (refer to Figure 5-10, page 198). The location of these crash types will be generally at intersections or driveways.

The intended purpose of the analysis by crash type is to analyze the relationships between crash mechanisms and injury severity at these hazardous locations. The proportions of severe injuries to total driver involvements are shown in Figure 5-13, page 216. For crashes occurring on high-speed multilane roads, head-on collisions are the most likely to result in severe injury, where more than 11% of the driver involvements result in severe injury. In a close second, in fixed object crashes more than 9% of the involvements resulted in a severe injury. In third and fourth place, the angle and left turn collisions had very similar proportions of severe injury involvements, around 8% of the total driver involvements in multilane arterial corridors. The rest of the crash types (right turn, sideswipe, rear-end and other) had about half the proportion of severe involvements of the angle and left turn collisions.

These results appear to follow a similar trend in regards to injury severity as that seen for other types of roads. Head on crashes comprised only 1.96% of the total and did not present an appropriate sample for this analysis. On the other hand, fixed object crashes are the most significant of the segment crashes in regards to their severity (see Figure 5-13, page 216). The analysis of fixed object crashes will allow a better understanding of the contributing factors for crashes in high-speed multilane roads to compare their contributing factors to those on limited access facilities and two lane roads.

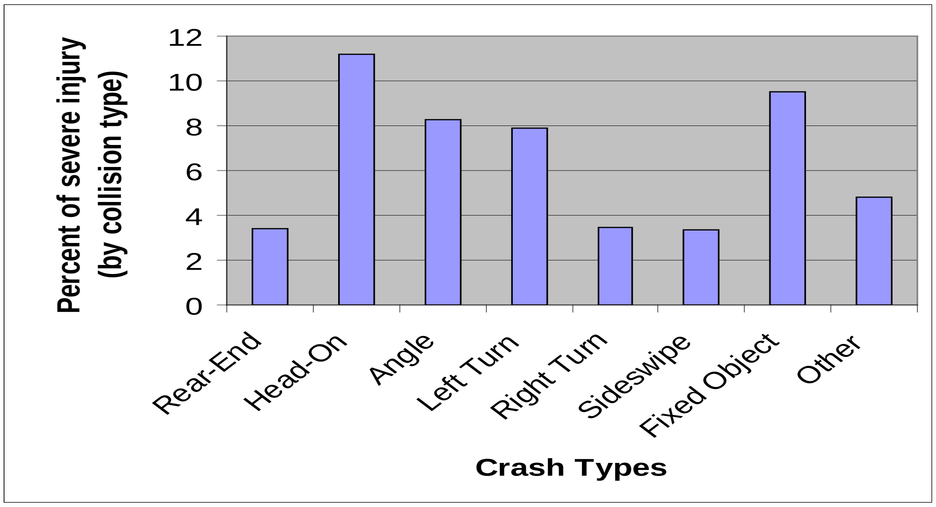


Figure 5-13: Severe Crashes Proportions by Crash Types on High-speed Multilane Roads

The distinctive factors affecting driver injury severity in rear-end, angle, left turn and fixed object crashes contributed to the goals of this investigation. It improved our understanding of the crash mechanisms and injury severity contributing factors. It also provided additional evidence towards a proposal for injury severity analysis for crashes on high-speed multilane roads. This discussion will focus on the differences between these models and the road entity models already discussed. Detailed coefficient values and standard errors for all of the models can be found in Appendix E.

5.4.3.1 Rear-end Crash Involvement Models

The most common crash type is the rear-end, but is also considered less likely to produce severe injury than the other types of crashes, except for the sideswipe collision. The models developed for rear-end crashes tend to be richer in covariates due to the higher frequencies of crashes, even with lower proportions of severe crashes. The unsignalized intersections exhibit higher proportions of severe injuries, as for the other crash type models except fixed object. The higher severity of crash involvements at or near unsignalized intersections is an expected result as shown in the analysis of Figure 5-13, page 216. The detailed model coefficients and standard errors are shown in Appendix E.

Table 5-50: Goodness of fit Measures for the Final Analysis Rear-end Crash Models

GOF Parameter	SIGNAL	UNSIG	SEGMENT	PURE SEG	OVERALL
Number of Variables	11	11	20	16	23
Degrees of freedom	17	19	34	27	42
Marginally significant coefficients	4	3	4	3	2
Non-significant coefficients	0	4	7	6	8
Sample size	10049	5701	20281	14580	33831
Response severe/non-severe ratio	3.24%	3.68%	3.54%	3.48%	3.43%
AIC	2656.988	1625.667	5500.224	3891.125	9084.313
Hosmer-Lemeshow p-value	0.9912	0.0277	0.0995	0.3388	0.4483
c value (area under ROC curve)	0.734	0.762	0.762	0.767	0.750
Percent Concordant	72.1	75.3	75.3	75.8	74.0
Adjusted R-squared	0.1009	0.1356	0.1416	0.1463	0.1240

All of the rear-end crash models attained acceptable goodness of fit, except for the unsignalized intersection model (see Table 5-50). While most of the goodness of fit measures improved for all of the models, the unsignalized intersection model failed the Hosmer-Lemeshow calibration tests. In addition, the adjusted R-squared values were lower for all of the rear-end crash models, which would traditionally be unexpected. It seems that the analysis of individual crash types yields less information in a model relative to the analysis at the corridor level or by road entities.

Table 5-51: Odds Ratios for Variables Found Significant in the Final Analysis Rear-end Crash Models

Variable	Level	Overall	Signal	Unsig	Segment	Pure Seg
Driver age 80-98 (vs. 25-64)	**5**	1.060†		0.678†	0.865†	0.914†
Driver age 65-79 (vs. 25-64)	**4**	1.228*	1.283*	1.631	1.084†	0.836†
Driver age 20-24 (vs. 25-64)	**3**	0.629	0.682*	0.422	0.585	0.620
Driver age 15-19 (vs. 25-64)	**2**	0.625	0.611	0.694*	0.630	0.586
Gender (Female vs. Male)		1.064†	0.663	0.681†	1.002†	1.647
Seat Belt Used (vs. no)	**1**	0.255	0.298	0.171	0.227	0.306
Gender*Seat Belt Used	**1**	1.590		3.259	1.908	
Speeding (Unknown vs. not)	**2**	0.632	0.585	0.748*	0.674	0.678
Speeding (Yes vs. Not)	**1**	0.409	0.311	0.512	0.460	0.426
At Fault driver (vs. not)		0.589	0.567		0.595	0.547
FL Resident (vs. not)		1.770	2.486			
Physical Defects (vs. not)		2.179	2.258		2.222	2.393
Driver Ejected (Yes/Partial vs. No)		1.817	4.427			
Point of Impact (driver side vs. not)				2.860		
Other vehicle type (vs. auto)	**5**	0.943†		1.219†	0.866†	0.983†
Bike/motorcycle (vs. auto)	**4**	2.319		2.962	3.645	4.635
Trucks/buses (vs. auto)	**3**	0.294		0.193	0.301	0.539*
Van/Light Truck/Pick up (vs. auto)	**2**	0.834		0.722*	0.850*	0.909†
Work Area (Entered vs. none)	**3**	0.771†			0.894†	
Work Area (Nearby vs. none)	**2**	0.530			0.450	
Private vehicle use (vs. not)						2.019
Concrete Surface (vs. other)		0.468			0.418	0.219
Urban area (vs. Rural)		0.812		0.587	0.791	
Speed limit (40-45 vs. other)	**1**	0.594	0.589	0.461	0.552	0.508
***adt* per Lane (thousands)**		0.947		0.940	0.929	0.926
Avg Truck Factor (percent)					1.019	1.032
Skid Res. (1≤FN<35 vs. FN≥35)		1.312			1.408	1.542
Lane<10 ft (vs. 11-12 ft)	**4**	0.752	0.674*			
10ft ≤lane< 11ft (vs. 11-12 ft)	**3**	0.659	0.465			
Lane > 12 ft (vs. 11-12 ft)	**2**	0.678	0.698*			
Full Non-High Mast (vs. none)	**Y**	1.359		1.909	1.455	1.326*
Partial Non-High Mast (vs. none)	**P**	0.755		0.850†	0.746	0.723*
High Masts (full vs. none)	**Y**	1.306†			1.605*	
High Masts (partial vs. none)	**P**	4.016			4.943*	
Non applicable (vs. class 2,3,4)	**9**	1.051†				
Access class 7 (vs. class 2,3,4)	**7**	0.767*				
Access class 6 (vs. class 2,3,4)	**6**	0.746				
Access class 5 (vs. class 2,3,4)	**5**	0.850				
FC-3, FC-6, N/A (vs. FC-2)	**9**	1.037†			1.060†	1.095†
Friction Course 5 (vs. FC-2)	**5**	0.797†			0.988†	1.064†
Friction Course 4 (vs. FC-2)	**4**	0.807			0.764	0.689
Friction Course 1 (vs. FC-2)	**1**	0.718			0.687	0.590
Sidewalk width ≥ 6 ft (vs. < 4 ft)	**3**	0.774	0.664		0.697	
4 ft ≤ Sidewalk < 6 ft (vs. < 4 ft)	**2**	0.827	0.584		0.860*	
Weekend (vs. Weekday)		0.760			0.717	0.718

Notes: * Effect is marginally significant (p<0.20); † Effect is not significant (p≥0.20)

Driver characteristics are not very significant in rear-end crash injury severity, as they were in the corridor models. Driver age loses significance in the rear-end models, especially for the older drivers. The very old driver level is no longer significant and the old driver level is only marginally significant. While there is no significant effect of the older drivers on rear-end crash severity, younger drivers have lower odds ratios in rear-end crashes than in all crashes (0.62 vs. 0.77). The gender variable is no longer significant, except for the signalized intersection and pure segment models. The gender variable here refers to males vs. females. While males are at an advantage in signalized intersections (OR=0.663), there are at a disadvantage in pure segments (OR=1.647). This may be driveway-related crashes or lane changing crashes with different mechanisms and exposure than those related to signalized intersections.

On the other hand, driver behavior continues to be important for the rear-end crash models, compared to the insignificance of driver characteristics. This agrees with an early study of the relationships between driver characteristics, behavior, vehicle types, crash types, and driver injury severity (Kim et al., 1995). The use of seat belts is more effective in rear-end crashes (OR=0.225) when compared to the overall model (all crashes, OR=0.303). This is also true for males (OR=1.590) vs. in the overall model (all crashes, OR=1.245). The speeding drivers had the same or very similar effects as in the overall model (all crashes, OR=0.409), as well as the at-fault drivers (0.589) and residence code (OR=1.770). Meanwhile, the ejected event has a negative, but much lower effect (OR=1.817) than in the overall model (all crashes, OR=4.270), while the effect is similar for the signalized intersection model (OR=4.427). Driver ejections from a rear-end crash with stopped vehicles at or near intersections generally involve high speed differentials, thus the highly negative effect. The point of impact was only significant in the unsignalized intersection model, but had a large negative effect (OR=2.860). Driver side impacts may come as a mismatch with rear-end crashes. However, there is at least one situation with the potential of rear-end crashes with impacts on the driver's side, as shown in Figure 5-14. In non-restricted median openings, vehicles waiting to make a turn are likely to encroach on the through travel lane increasing the risk for a rear-end crash with a stopped vehicle on the drivers side (back), which tends to be severe.

Figure 5-14: Potential Rear-end Crash with Point of Impact on Driver's Side on High-speed Multilane Roads (Source: Microsoft Virtual Earth)

The type of vehicle continues to be an important variable for the rear-end crash models. Two main differences are evident for the heavy truck drivers and the bicyclists and motorcyclists. The drivers of heavy trucks have an increased positive effect (OR=0.294 vs. 0.357 in overall model-all crashes), all else held constant. This suggests that in rear-end crashes, passenger vehicle drivers are at a major disadvantage with relation with heavy trucks. In addition, the bicycle and motorcycle vehicles are now at a major disadvantage with respect to passenger cars (OR=2.319) for rear-end crashes. Motorcyclist maneuvers, including improper lane changing (overtaking) may have a role in this situation. Also, the motor vehicle interactions with bicycles in the rightmost lane (with wide lane or designated bike lane) are likely to play a role in this situation, since this variable was significant in the segment models, but not in the signalized intersection model.

Crashes occurring in the work zone transition (entering) were significant in the injury severity models for rear-end crashes. In the case of rear-end crashes, the positive effect was increased (OR=0.530), which is expected for the type of collision that tends to be less severe than other types.

In terms of roadway characteristics, the concrete surface factor proved to be significant for rear-end crashes in segments and in the overall model (rear-end crashes). Concrete surfaces have a positive effects with respect to driver injury severity (OR=0.468) when compared to other pavement surfaces. These concrete surfaces of the high-speed multilane roads tend to be in downtown areas and the relationship with land use is possible in this case because land use affects operating speeds and traffic signal metering, among other important factors affecting traffic flow. The urban areas have a slightly higher positive effect for the rear-end crashes (OR=0.812). It is possible than lower operating speeds are more likely to affect the severity of the rear-end crashes than other types. Indeed lower operating speeds (40-45 mph speed limit) have an increased positive effect on driver injury severity (holding all else constant) for rear-end crashes (OR=0.594) than in the overall model (all crashes, OR=0.676). The positive effect of traffic volumes (OR=0.947) is slightly increased, while the truck volume factor looses significance in the rear-end crash model. The skid resistance effect is greater (OR=1.312) holding all else constant, but not in the presence of additional interactions. The presence of the concrete surface variable points to a possible increased importance of pavement characteristics in the rear-end crash injury severity models.

The lane width effects favored the widths between 11 and 12 ft, which is similar to the effect in the alternative urban area model. Intuition suggests that this relationship is likely to change by land use. In the two lighting variables the differences between the partial and full lighting were increased in comparison to the overall model (all crashes). This suggests that full roadway lighting is more beneficial in reducing injury severity of rear-end crashes. Likewise, the benefits of access management, especially class 6 (OR=0.746) were increased in comparison to the overall model (all crashes, OR=0.833). A similar trend followed with the friction course and sidewalk width. The increased benefits of improved roadway characteristics are most notable in the rear-end crashes, which are the most frequent. A final note is that drivers traveling on weekends have an increased positive effect (OR=0.760) vs. the overall model (all crashes, OR=0.906).

5.4.3.2 Angle Crash Involvement Models

The angle crash injury severity models generally had acceptable goodness of fit measures except for the combination segment and no-signalized intersection model, which failed the Hosmer-Lemeshow test (see Table 5-52, page 223). This evidence of misspecification in the overall angle crash model is likely due to the land use differences, as shown in the variables to be discussed. However, the adjusted R-squared values denoted improved realized model information potential over the rear-end model. Many of the variables loose significance; the interpretative power of the models were degraded. The unsignalized intersections exhibit higher proportions of severe injuries, as for the rest of the crash types models correlated to intersections (rear-end, angle and left turn), as expected. The model coefficients and standard errors are detailed in Appendix E.

Table 5-52: Goodness of fit Measures for the Final Analysis Angle Crash Models

GOF Parameter	SIGNAL	UNSIG	SEGMENT	PURE SEG	OVERALL
Number of Variables	11	12	15	12	19
Degrees of freedom	20	18	29	23	35
Marginally significant coefficients	0	0	1	2	2
Non-significant coefficients	4	2	7	7	6
Sample size	5401	6273	11052	4779	16802
Response severe/non-severe ratio	7.55%	9.58%	8.51%	7.11%	8.10%
AIC	2631.097	3434.963	5605.323	2169.483	8323.191
Hosmer-Lemeshow p-value	0.7879	0.2267	0.0505	0.8162	0.0483
c value (area under ROC curve)	0.728	0.750	0.748	0.747	0.741
Percent Concordant	72.3	74.6	74.4	74.0	73.7
Adjusted R-squared	0.1316	0.1839	0.1754	0.1667	0.1601

The amount of driver vehicle- and crash-related variables in this model remained relatively stable, with the addition of the contributing cause and vision obstructed (see Table 5-53, page 224). In regards to roadway-related variables, auxiliary lanes variables and friction course were not significant in these models, but the shoulder width entered the overall angle crash model. The vision obstructed variable was significant only in the segment models, making it a new variable in comparison to the overall and rear-end models. Meanwhile, the contributing cause was also significant in the unsignalized intersection models. This was expected as angle crashes are likely to be influenced by visibility problems as well as driver impatience (aggressiveness) and impairment (DUI) in driveways and minor roads.

The effects in driver age group had an inversed situation when compared to the rear-end crashes. Old and very old drivers had a significant disadvantage (OR=1.663 and 2.099) when compared to middle age drivers in angle crashes. Gender was also significant, with the females at a significant disadvantage with respect to male drivers (OR=1.446), higher than in the overall model (all crashes, OR=1.217). The point of impact is on the side of the vehicle, affording less protection to the driver, increasing the importance of physiological conditions on crash outcome.

Table 5-53: Odds Ratios for Variables Found Significant in the Final Analysis Angle Crash Models

Variable	Level	Overall	Signal	Unsig	Segment	Pure
Driver age 80-98 (vs. 25-64)	**5**	2.099	2.484	2.638	2.127	1.494†
Driver age 65-79 (vs. 25-64)	**4**	1.663	1.443	1.991	1.887	1.886
Driver age 20-24 (vs. 25-64)	**3**	0.852*	0.666	1.057†	0.962†	0.843†
Driver age 15-19 (vs. 25-64)	**2**	0.943†	0.983†	1.034†	0.938†	0.816†
Gender (Female vs. Male)		1.446	1.678	1.525	1.404	1.318
Seat Belt Used (vs. no)	**1**	0.301	0.308	0.305	0.303	0.294
Speeding (Unknown vs. not)	**2**	0.693	0.794†	0.454	0.625	0.869†
Speeding (Yes vs. Not)	**1**	0.640	0.690	0.607	0.593	0.536
Other Cont. Cause (vs. none)	**4**	1.482		1.774	1.890	1.849
Aggressive Driving (vs. no improper)	**3**	2.566		5.960	3.904	2.333*
Alcohol/Drug use (vs. no improper act)	**2**	1.778		2.452	2.100	1.572
At Fault driver (vs. not)		0.469		0.310	0.391	0.521
Red light running (vs. not)		1.326				
Physical Defects (vs. not)		1.549				
Driver Ejected (Yes/Partial vs. No)		4.163	5.443	5.312	4.127	3.659
Off Roadway (vs. not)		1.577			1.773	
Impact (driver side vs. not)		1.791	1.878	2.078	1.846	1.482
Other vehicle type (vs. auto)	**5**	0.684†			0.816†	0.733†
Bike/motorcycle (vs. auto)	**4**	1.097†			1.043†	1.042†
Trucks/buses (vs. auto)	**3**	0.254			0.260	0.230
Van/Light Truck/Pick up (vs. auto)	**2**	0.793			0.750	0.726
Private vehicle use (vs. not)				2.019		
Urban area (vs. Rural)		0.689	0.733	0.570	0.635	0.666
Intersection (vs. not)		1.315			1.396	
Speed limit (40-45 vs. other)	**1**	0.683	0.638	0.607	0.677	0.742
***adt* per Lane (thousands)**		0.969				
Skid Res. (1≤FN<35 vs. FN≥35)				1.240		
Lane<10 ft (vs. 11-12 ft)	**4**	0.737				
10ft ≤lane< 11ft (vs. 11-12 ft)	**3**	0.835†				
Lane > 12 ft (vs. 11-12 ft)	**2**	0.822*				
Curb Shoulder (vs. Paved)	**3**		1.660			
Unpaved Shoulder (vs. Paved)	**2**		0.949†			
Shoulder width ≥ 10 ft (vs. <6ft)	**4**	0.837†				
8 ft ≤ Shoulder width < 10 ft (vs. <6ft)	**3**	0.649				
6 ft ≤ Shoulder width < 8 ft (vs. <6ft)	**2**	0.633				
Non applicable (vs. class 2,3,4)	**9**	0.906†	0.879†		0.891†	
Access class 7 (vs. class 2,3,4)	**7**	0.623	0.578		0.619	
Access class 6 (vs. class 2,3,4)	**6**	0.695	0.361		0.842†	
Access class 5 (vs. class 2,3,4)	**5**	0.756	0.698		0.752	
Sidewalk width ≥ 6 ft (vs. < 4 ft)	**3**		0.579			
4 ft ≤ Sidewalk < 6 ft (vs. < 4 ft)	**2**		0.720			
Other (vs. no vision obstruction)	**4**				0.835†	0.523†
Parked/Stopped Veh (vs. no obstruction)	**3**				1.227*	1.488*
Bad Weather, Smoke, Glare (vs. no obst)	**2**				1.749	3.555

Notes: * Effect is marginally significant (p<0.20); †

The physical defects and ejection events had similar effects than in the overall model (all crashes). Driver behavior was important in the angle crash models. Use of seat belts remained similar to the effect on the overall model (all crashes). Likewise, drivers at fault have a similar positive effect than in the overall model (all crashes). It would be of interest to investigate the relationship between at fault drivers and collision sequence to find out whether these are correlated with the striking vehicle. The harmful effects of speeding, aggressive driving, driving under the influence and red light running were increased with respect to the overall model (all crashes). These results suggest a strong correlation between driving behavior and angle crashes.

The results with respect to the crash- and vehicle-related variables suggest that vehicle protection becomes more important in angle crashes and off road (driveway) crashes are an important contributing factor. The off roadway negative effect (OR=1.577) contrasts sharply with the positive effect in the overall model (all crashes, OR=0.613). An off road angle crash will generally be related to driveways. Access management, which includes driveway location and design, would play an important role in preventing some of these highly severe crashes. Vehicle protection and configuration is important in the angle crash outcomes. Drivers in angle crashes with vehicle impact on their side had a significant negative effect (OR=1.791) vs. a non-significant effect in the overall model (all crashes). Both the type of vehicle and private vehicle use exhibited an increased negative effect for the passenger car drivers in angle crashes when compared to the overall model (all crashes).

The speed limit and *adt* per lane variables retained similar effects when compared to the overall model (all crashes). Meanwhile the injury severity difference between involvements in urban and rural areas was more acute in the angle crash models (OR= 0.689) when compared to the overall model (all crashes, OR=0.879). Skid resistance lost significance in the overall model (angle crashes), but the effect in the unsignalized intersection model remained similar to that in the respective (unsig) model for all crashes. The lane width effect for lanes less than 12 ft was not significant. On the other hand, the positive effect of lanes wider than 12 ft (OR=0.737) was increased for angle crashes in comparison to the overall model (all crashes, OR=0.827). It would be counterintuitive that increased crossing distance for the road would reduce angle crash frequency or severity. However, two factors have to be taken into account: congestion in urban areas with multilane roads of more than four lanes and the rightmost lane width effect on increased visibility for the driver on a minor road (or driveway). In urban areas with visibility obstructions (generally with very narrow curb shoulder), a wider rightmost lane allows the driver a better view of the coming traffic. The type of shoulder variable was only significant for one coefficient in the signalized intersection model, with a negative effect of curb shoulders (OR=1.660), higher than in the overall model (all crashes, OR=1.162). On the other hand, increased shoulder width from less than 3.5 ft (OR=0.633) to between 3.5 and 6 ft (OR=0.649) only had a small increase in the positive benefit. The relationships between land use, lane width, shoulder type and width; and injury severity in corridors need further investigation.

The positive effects of increased access management are more noticeable in the angle crash models, with class 5 (OR=0.756) exhibiting the most benefit vs. the overall model (all crashes, OR=0.879). Even tough class 5-7 have similar intersection spacing, the medians openings are restricted for class 5. The positive effect of the median opening restrictions is suggested for one of the most severe collision types. There is an increased benefit for the sidewalks greater than 6 ft in angle crashes at or near signalized intersections (OR=0.579) when compared to all collision types (OR=0.665); this denotes the important role of increased visibility in reducing angle crash severity.

5.4.3.3 Left turn Crash Involvement Models

The left turn models exhibited acceptable goodness of fit (see Table 5-54). The detailed model coefficients and standard errors are shown in Appendix E. In terms of the significant coefficients, it performed very well in comparison to the overall model (all crashes). The AIC values are generally lower than for the rear-end and angle crash models, while the adjusted R-squared are lower than the overall model (all crashes), similar to the other crash type models. The unsignalized intersections exhibit higher proportions of severe injuries, expected.

Table 5-54: Goodness of fit Measures for the Final Analysis Left turn Crash Models

GOF Parameter	SIGNAL	UNSIG	SEGMENT	PURE SEG	OVERALL
Number of Variables	8	14	10	6	14
Degrees of freedom	17	25	13	9	24
Marginally significant coefficients	1	4	0	0	4
Non-significant coefficients	6	4	2	1	4
Sample size	2941	6289	3993	1388	7023
Response severe/non-severe ratio	8.77%	9.59%	7.54%	6.34%	7.99%
AIC	1623.575	3433.337	1965.075	605.252	3589.096
Hosmer-Lemeshow p-value	0.7344	0.7437	0.8523	0.6008	0.2401
c value (area under ROC curve)	0.718	0.759	0.710	0.743	0.722
Percent Concordant	71.0	75.5	70.1	71.8	71.7
Adjusted R-squared	0.1187	0.1920	0.1168	0.1315	0.1209

The variables found significant in the left turn models present a combination of driver-, vehicle-, crash-, road- and environment-related variables similar to that in other crash type models (see Table 5-55, page 228). The common feature of these models is that most changes occur in crash and roadway variables, while most driver and vehicle variables stayed in the models. The lack of many variables in the overall (left turn crashes) model and the similar effects of the variables to the angle crash model allows a comparison between the two crash types (in terms of injury severity) that may justify additional analysis in future research.

Table 5-55: Odds Ratios for Variables Found Significant in the Final Analysis Left turn Crash Models

Variable	Level	Overall	Signal	Unsig	Segment	Pure Seg
Driver age 80-98 (vs. 25-64)	**5**	2.099	1.798	2.623	2.467	
Driver age 65-79 (vs. 25-64)	**4**	1.588	1.546	1.979	1.754	
Driver age 20-24 (vs. 25-64)	**3**	0.855†	0.902†	1.049†	0.841†	
Driver age 15-19 (vs. 25-64)	**2**	0.757*	0.745†	1.022†	0.826†	
Gender (Female vs. Male)		1.542	1.585	1.528	1.661	2.543
Seat Belt Used (vs. no)	**1**	0.305	0.310	0.306	0.309	0.374
Speeding (Unknown vs. not)	**2**			0.451		
Speeding (Yes vs. Not)	**1**			0.611		
Other Cont. Cause (vs. none)	**4**			1.849		
Aggressive Driving (vs. no improper)	**3**			5.996		
Alcohol/Drug use (vs. no improper act)	**2**			2.466		
At Fault driver (vs. not)		0.578	0.534	0.308	0.615	
Driver Ejected (Yes/Partial vs. No)		5.487	6.746	5.390	4.031	8.869
Off Roadway (vs. not)		3.674			3.855	
Impact (driver side vs. not)		1.529		2.111	1.738	2.257
Other vehicle type (vs. auto)	**5**	0.327*	0.000†			
Bike/motorcycle (vs. auto)	**4**	0.572*	0.476†			
Trucks/buses (vs. auto)	**3**	0.316	0.133			
Van/Light Truck/Pick up (vs. auto)	**2**	0.650	0.656			
Private vehicle use (vs. not)				2.012		
Urban area (vs. Rural)				0.566	0.714	
Median Type (paved vs. raised)	**2**	0.697			0.547	
Speed limit (40-45 vs. other)	**1**	0.653	0.664	0.623	0.586	
Avg Truck Factor (percent)		1.029				
Skid Res. (1≤FN<35 vs. FN≥35)		1.227		1.277		
Full Non-High Mast (vs. none)	**Y**	1.515				
Partial Non-High Mast (vs. none)	**P**	0.848†				
Non applicable (vs. class 2,3,4)	**9**	1.248*	1.482*			0.949†
Access class 7 (vs. class 2,3,4)	**7**	0.433	0.222			0.175
Access class 6 (vs. class 2,3,4)	**6**	0.918†	1.027†			0.403
Access class 5 (vs. class 2,3,4)	**5**	0.881†	0.869†			0.545
3 or more Auxiliary Lanes (vs. none)	**3**			1.204†		
2 Auxiliary Lanes (vs. none)	**2**			0.691		
1 Auxiliary Lanes (vs. none)	**1**			1.215*		
FC-3, FC-6, N/A (vs. FC-2)	**9**			1.219*		
Friction Course 5 (vs. FC-2)	**5**			1.505*		
Friction Course 4 (vs. FC-2)	**4**			1.019†		
Friction Course 1 (vs. FC-2)	**1**			0.697*		
Weekend (vs. Weekday)						1.764

Notes: * Effect is marginally significant (p<0.20); † Effect is not significant

The driver-related variables had generally similar effects than the angle crash models. The effects in driver age group had an inversed situation when compared to the rear-end crashes. Old and very old drivers had a significant disadvantage (OR=1.588 and 2.099) when compared to middle age drivers in left turn crashes, almost equal to the effects in angle crashes. Gender was also significant, with the females at a significant disadvantage with respect to male drivers (OR=1.542), higher than in the overall model (all crashes, OR=1.217). In these types of crashes, the point of impact is generally on the side of the vehicle, affording less protection to the driver, thus the increased importance of physiological conditions in the crash severity outcome. The physical defects factor was no longer significant in the left turn crash model. Meanwhile the ejection events had an increased negative effect (OR=5.487) than in the overall model (all crashes, OR=4.270). Driver behavior was not as important in the left turn crashes as in the angle crash models. The effect of seat belt usage and drivers at fault remained similar to the corresponding effects on the overall model (all crashes). However, the harmful effects of speeding, aggressive driving, driving under the influence and red light running were only significant for the unsignalized intersections. For left turn and angle crashes at or near unsignalized intersections, the effects (odds ratios) of speeding, aggressive driving and driving under the influence were similar to those in the angle crash models. These results suggest a strong correlation between driving behavior and left turn crashes for the unsignalized intersections and a similarity with left turn crashes, with respect to driver injury severity.

Meanwhile, the crash and vehicle factors had similar effects that in the angle crash models. The vehicle type and crash configuration effects were also increased in relation to the overall model (all crashes). The off roadway negative effect (OR=3.674, compared to OR=1.577 for angle crashes) contrasts sharply with both the positive effect in the angle and overall models (all crashes, OR=0.613). The results suggest similar crash mechanisms of angle and left turn crashes and that off road (driveway) crashes are an increasingly important contributing factor, more so than in angle crashes. An off road angle crash will generally be related to driveways. Access management, which includes driveway location and design, would play an important role in preventing some of these highly severe crashes. Vehicle protection and configuration is important in the left crash outcomes. Drivers in left turn crashes with vehicle impact on their side had a significant negative effect (OR=1.529) vs. a non-significant effect in the overall model (all crashes). Both the type of vehicle and private vehicle use exhibited an increased negative effect for the passenger car drivers in left crashes when compared to the angle and overall model (all crashes). Compared to angle crashes, the positive effects for LTV drivers was increased (OR=0.650 vs. 0.793 in angle crashes), while for the heavy trucks it was decreased (OR=0.316 vs. 0.254 in angle crashes). This suggests an increased negative effect on the injury severity of drivers of passenger vehicles for left turn crashes.

In terms of roadway-related variables, the important variable land use was not significant for the left turn crash overall model. The good performance of this model and the lack of significance of the land use add to the proof of the important effect of land use on the slight misspecification in the overall angle crash model. In unsignalized intersections in urban areas, the positive effect on injury severity was increased (OR=0.566) when comparing to the overall model (all crashes, OR=0.808), suggesting an increasing negative effect of unsignalized intersections in rural areas in the severity of left turn crashes. The median type paved two-way turn lane (TWTL) significance in the segment model suggests a positive effect of TWTL (OR=0.697) vs. raised medians for the left turn crashes in midblock locations. Although TWTL may increase the chance of head-on collisions, it is also true that provides a *refuge* space to allow midblock left turn movements from the minor roads into the multilane facility in two stages. It also improves the chances of avoiding oncoming traffic flows in both directions, in contrast to the example shown in Figure 5-14, page 220, where the median opening does not provide enough space for a left turn from the major road. These median treatments are applied for different traffic and road (drainage) conditions.

The speed limit and average truck factor variables retained similar effects when compared to the overall model (all crashes). The *adt* per lane was not significant in the left turn crash model. The driver injury severity in left turn crashes was negatively influenced by increasing heavy truck traffic (OR=1.029). This effect was not significant in the angle crash models. This might be related to the different acceleration characteristics for a left turn of heavy truck. In addition, the heavy truck size (and mass) may be a contributing factor increasing the severity of left turn crashes. The non significance of *adt* agrees with recent research modeling injury severity for total left turn crashes related to signalized intersections in arterial corridors which also showed traffic volume significance for certain patterns of left turn crashes (Wang and Abdel-Aty, 2008). Additional research is needed to compare left turn crashes at unsignalized and driveway-related midblock locations to those in signalized intersections in high-speed multilane arterials.

Increasing skid resistance had increased negative effect on left turn crashes (OR=1.227) compared to the overall model (all crashes, OR=1.198). For the overall angle crash model this effect was not significant. In addition, the effect in the unsignalized intersection model (OR=1.277) was higher than the corresponding effect in the model (unsig) for angle crashes (OR=1.240). The increased importance of skid resistance for crashes at intersections involving a left turn movement was expected. However, the type of friction course was not significant for the (overall) left turn crashes, but only significant for the unsignalized intersection model. The through lane width effect was not significant for left turn crashes, which was also as expected. On the other hand, the number of auxiliary lanes did have a significant positive effect. Drivers making left turns from a double left turn lane unsignalized approach on the major road had a positive effect on injury severity (OR=0.691), larger than in the overall mode (all crashes, OR=0.882). This effect might be due to the presence of double left turn lanes on closely spaced intersections in urban areas.

The positive effects of increased access management on driver injury severity are not manifest in the left turn crashes as they are for the angle crashes. Only class 7 is significant for left turn crashes (OR=0.433) with an increased positive benefit when compared to angle crashes (OR=0.623). This is probably due to the positive effect of traffic volumes in urban closely spaced intersections (class 7) on left turn crash severity. Traffic volumes and land use may have an important effect on the left turn crashes during weekends, which has a negative effect (OR=1.764) in contrast with the positive effect for the overall model (all crashes, OR=0.906). It seems that there are indirect influences of traffic volume and land use in driver injury severity outcomes of left turn crashes.

5.4.3.4 Fixed Object Crash Involvement Models

Fixed object crashes are the most severe crash types in this analysis. There was sufficient numbers of these crashes to develop a set of models with acceptable goodness of fit measures (see Table 5-56, page 233). The detailed model coefficients and standard errors are shown in Appendix E.

Table 5-56: Goodness of fit Measures for the Final Analysis of Fixed Object Crash Models

GOF Parameter	SIGNAL	UNSIG	SEGMENT	PURE SEG	OVERALL
Number of Variables	2	7	14	13	7
Degrees of freedom	3	10	23	21	12
Marginally significant coefficients	0	0	3	3	2
Non-significant coefficients	1	3	4	4	1
Sample size	568	761	4298	3537	5011
Response severe/non-severe ratio	12.32%	9.59%	14.61%	15.69%	14.29%
AIC	412.638	437.989	3042.54	2619.059	3597.015
Hosmer-Lemeshow p-value	0.7026	0.8551	0.7322	0.2587	0.1100
c value (area under ROC curve)	0.642	0.754	0.767	0.770	0.743
Percent Concordant	51.7	74.2	76.5	76.8	74.0
Adjusted R-squared	0.0640	0.1748	0.2238	0.2265	0.1824

The variables entered into these models were considerably less than in the previous models (see Table 5-57, page 234). The different crash mechanism and the presumed prominence of single vehicle crashes lead to fewer interactions with other vehicles and other road. However, driver characteristics and behavior can be important factors for these crashes. Another main difference is that the most effective model is the segment combined with unsignalized intersections. This is expected as most fixed object crashes occur on road segments. A notable exception is unsignalized intersections on a roadway curve, which showed a significant negative effect.

Some important differences with the road entity models were found in the fixed object crash model. Traditional conceptions of off roadway crashes on two-lane roads are mostly not applicable to high-speed multilane arterial corridors. This analysis is important since it shows some of the road characteristics that affect crashes that generally occur outside of the travel lanes. The precursors to these crashes may be varied, but the outcome of the crash is more likely to be influenced by roadside objects or other road features than other types of crashes.

Table 5-57: Odds Ratios for Variables Found Significant in the Final Analysis of Fixed Object Crash Models

Variable	Level	Segment	Pure Segment	Unsig	Overall	Signal
Driver age 65-98 (vs. 25-64)	4	1.507	1.145†	3.507		
Driver age 20-24 (vs. 25-64)	3	0.847*	0.801*	1.063†		
Driver age 15-19 (vs. 25-64)	2	0.733	0.699	0.797†		
Driver Ejected (Yes/Partial vs. No)		12.511	4.119	5.706	3.945	
Female*Ejected		3.794				
Gender (Female vs. Male)		1.240		2.138		
Seat Belt Used (vs. no)	1	0.368	0.378	0.379	0.395	
Speeding (Unknown vs. not)	2	1.061†	1.103†	0.753†	1.184†	2.627
Speeding (Yes vs. Not)	1	0.530	0.589	0.354	0.582	0.835†
Other Cont. Cause (vs. none)	4	2.599	2.190		2.327	
Aggressive Driving (vs. no improper)	3	2.291	1.869		1.768	
Alcohol/Drug use (vs. no improper act)	2	2.050	1.937		1.869	
At Fault driver (vs. not)		0.467	0.443		0.446	
Off Roadway (vs. not)				1.803		
Impact (driver side vs. not)						2.850
Other maneuver (vs. straight ahead)	4	1.254*	1.380*		1.238*	
Left Turn (vs. straight ahead)	3	0.341	0.439		0.434	
Slowing / Stopping (vs. straight ahead)	2	1.284†	1.470†		1.588*	
Private vehicle use (vs. not)			2.245			
Multivehicle (vs. single vehicle crash)			0.665			
Urban area (vs. Rural)		0.711	0.740		0.914	
Roadway Curve (vs. not)		1.620	1.481	2.625		
Curb Shoulder (vs. Paved)	3	1.094†	1.135†			
Unpaved Shoulder (vs. Paved)	2	0.802*	0.821*			
***adt* per Lane (thousands)**		0.909	0.904			
Full Non-High Mast (vs. none)	Y	1.647				
Partial Non-High Mast (vs. none)	P	1.047†				

Notes: * Effect is marginally significant (p<0.20); † Effect is not significant (p≥0.20)

The driver age variable was reduced to four groups were old and very old drivers were combined into one category. Older drivers 65 and more experienced a significant negative effect (OR=1.507), while young drivers had no significant effect and very young drivers had a positive effect (OR=0.733). Meanwhile the gender (female vs. male drivers) had also a significant effect (OR=1.240) which suggests that driver characteristics affect injury severity in fixed object crashes. The positive effect of seat belt usage is increased (OR=0.368) in comparison to the overall model (all crashes, OR=0.303). The ejection event had the greatest magnitude (12.511) of all odds ratios in the models. There was also a significant relationship between gender and the ejection event (OR=3.794), which is an additional negative effect for female drivers in fixed object crashes. The negative effects of unsafe driver behaviors including speeding (OR=0.530), driving under the influence (OR=2.291) and aggressive driving (2.050) were increased when compared to the overall model (all crashes, OR's=0.409, 1.748 and 1.593, respectively). On the other hand, drivers at fault had a positive effect (OR=0.467). This suggests that in the fixed object crashes, most driver-related effects were amplified.

A few crash variables were significant in this model. Interestingly enough, type of vehicle was not significant. For fixed object crashes at or near signalized intersections, there is a negative effect (OR=2.850) on injury severity for drivers that hit an object on their side of the vehicle. For drivers on left turn maneuvers, there is a positive effect on injury severity (OR=0.341) significant on the segment model that suggests a correlation with driveway crashes. It seems that these fixed object crashes refer to drivers on a left turn from the major road into a driveway with a fixed object (such as a sign), which are not as likely to result in severe injury as other fixed object crashes. Drivers of private vehicles (usually passenger cars) are at a disadvantage (OR=2.245) in terms of injury severity when involved in a fixed object crash in a high-speed multilane road segment. However, drivers involved in multivehicle crashes and with a fixed object are at an advantage (OR=0.665) over those involved in single vehicle crashes.

The off roadway crash had a negative effect for the unsignalized intersections (OR=1.803), which is contrary to the effect on all crashes (OR=0.613). This point to an area where certain characteristics tends to change the severity risk for fixed object crashes. There was a highly negative effect for fixed object crashes on unsignalized intersections on roadway curves (OR=2.625) on driver injury severity, much higher than for the overall model (all crashes, OR=1.510). This might be due to the tendency of utility poles to cluster around unsignalized intersections (by necessity) and limited right of way or poor utility planning. One example of a utility pole placed too close to the intersection corner in a curved road section is displayed in Figure 5-15.

Figure 5-15: Examples of Fixed Object (Utility Pole) Close to the Corner of an Unsignalized Intersection in a Curved Section of SR-423 (Lee Road) (Source: Microsoft Virtual Earth)

The positive effect of urban areas on driver injury severity after a fixed object crash (OR=0.711) is greater when compared to all crashes (OR=0.879). The negative effect of roadway tangent sections is significant (OR=1.620) and more pronounced than in the overall model (all crashes, OR=1.306). The type of shoulder is only marginally significant but it suggests a benefit of unpaved shoulders vs. a curb shoulder. It is likely that there is collinearity with land use. An important effect is traffic volume, with pronounced benefit OR=0.909) in comparison with the overall model (all crashes, OR=0.972). Another important road design characteristic that was significant, lighting density, exhibited a negative effect (OR=1.647) more so than in the overall model (all crashes, OR=1.129). This result has to be interpreted in the context of land use and a possible offset of the road lighting benefit by driver unsafe behavior, such as speeding. The relationships between these factors in the fixed object crash models and land use should be further investigated.

5.4.3 Relative Variable Significance

The relative variable significance method described in Section 3.5.4 and applied in Section 4.2.2 in the exploratory analysis was analyzed for the road entity models. In addition, an analysis of the relative significance of the variables in the crash type's models is presented. The variable relative significance of the best performing models for rear-end, angle, left turn and the fixed object crashes are presented here. The analysis of variable significance will allow assessing the most important effects on driver injury severity. There are other unobserved effects likely to affect the crash outcomes. However, by analyzing the most important effects with the available data, some unobserved factor may be intuitively theorized

In general, there are many factors affecting crash occurrence that cannot be obtained in a police crash report. Our main source of knowledge of crash occurrence and involvements comes from these reports which have known limitations. Within these limitations, however, it is possible to obtain valuable information and make conclusions about crash occurrence. It is likely that conditional crash outcomes (given that a crash occurs) could be more specifically predicted because there is more post crash event information available.

Table 5-58: Variable Relative Significance in the Final Analysis Road Entity Models

Variable	Overall	Inters	Signal	Segment	Pure Segment	Non-signal
Driver_Ageg_Group_x	Major	Major	Major	Major	Major	Major
Ejected_x	Major	Major	Major	Major	Major	Major
Speeding_x	Major	Major	Major	Major	Major	Major
Gender_x	Moderate	Major	Major	Minor	Major	Major
Safety_Equipment_x	Major	Major	Major	Major	Major	Major
Gender_x*Safety_Equipment	Moderate			Moderate		
At_Fault_driver_x	Major	Major	Major	Major	Major	Major
Red_light_running_x		Moderate	Moderate			
Residence_Code_x	Minor	Moderate	Moderate			
Physical_Defects_x	Major	Moderate	Moderate	Major	Major	
Harmful_Event_Group_	Major	Major	Major	Major	Major	Major
Contributing_Cause_x	Major	Major		Major	Major	Major
Type_of_Vehicle_x	Major	Major	Major	Major	Major	Major
Vehicle_Maneuver_x			Moderate			
Point_Impact_x	Marginal	Marginal	Major	Marginal		Major
Point_Impact*Speeding_x	Moderate	Moderate		Minor		
Off_Roadway	Major			Major	Major	
Off_Roadway*Speeding_x	Moderate			Major	Moderate	
Off_Roadway*Multivehicle	Major					
nWork_Area_x	Moderate			Moderate		
Multivehicle	Major					
Intersection*Multivehicle	Major					
Private_vehicle_use_				Moderate		
Private_veh*nAVGTFACT				Minor		
Speed_limit_x	Major	Major	Major	Major	Major	Major
ADT_PER_LANE	Major			Major	Major	Moderate
nAVGTFACT	Moderate			Moderate	Moderate	
LIGHTING	Moderate			Minor		
Traffic_Control	Minor			Major		
Access_class	Major	Major	Major			
nRural_Urban	Minor	Major	Moderate	Moderate	Marginal	Major
nType_of_Shoulder	Moderate	Minor	Minor			Moderate
Lane_width	Major	Major	Moderate	Moderate		
Roadway_Curve	Major			Major	Moderate	
Sidewalk_width_group	Major	Major	Major	Moderate	Minor	Moderate
LIGHTCDE	Major	Minor		Moderate	Moderate	
Type_Friction_Course	Major	Moderate	Moderate	Moderate	Moderate	
Intersection	Marginal					
Intersection*nRural_Urb	Moderate					
Skid_Resistance	Major	Minor		Major	Major	
nRural_Urban*Skid_Resis	Marginal	Marginal		Marginal	Minor	
AUX_Lane_Num				Moderate		Moderate
Day_of_Week	Moderate					

Note: Major ($p<0.001$), Moderate ($0.001 \leq p<0.01$) , Minor ($0.01 \leq p \leq 0.05$) and Marginal ($p>0.05$)

In the road entity models, most driver-, crash- and vehicle-related variables have a major significance (see Table 5-58, page 238). Driver age, speeding, driver at fault, contributing cause and physical defects had a major relative importance. Gender and the gender interaction with safety equipment had a moderate importance. Meanwhile, red light running only had moderate importance in the signal intersection model. Among the crash and vehicle variables, most variables had major relative importance. These included ejected event, harmful event (collision type), type of vehicle, off roadway (segment models), multivehicle and point of impact. Other variables such as work area and some off roadway and point of impact interactions had moderate importance. In the case of roadway variables, the major relative significance variables (in the overall model) included speed limit, *adt* per lane (segment models), access class (signalized intersections), lane width, roadway curve (segment models), sidewalk width, lighting (non-high mast), skid resistance (segment models) and type friction course. Other variables with moderate relative significance were average truck factors, type of shoulder and high-mast lighting.

One of the major differences between these final models and the models in the exploratory analysis was the variable relative significance. Few of the roadway-related variables available in the crash report had a major importance in the exploratory models. Only speed limit, *adt* per lane, and traffic control had major importance in some of the earlier models. On the other hand, many roadway-related characteristics had major importance in the final analysis. This is due in part to the better data preparation without data repetition. Also, the addition of a significant amount of roadway information from RCI was pivotal in improving the final analysis models and the results of this investigation.

Table 5-59: Variable Relative Significance in the Final Analysis Crash Type Models

Variable	Rear-End	Angle	Left Turn	Fixed Object
	Segment	Unsig	Unsig	Segment
Driver_Ageg_Group_x	Major	Major	Major	Moderate
Ejected_x		Major	Major	Major
Speeding_x	Major	Major	Major	Major
Gender_x	Marginal	Major	Major	Minor
Gender_x*Ejected_x				Moderate
Safety_Equipment_x	Major	Major	Major	Major
Gender_x*Safety_Equipment	Moderate			
At_Fault_driver_x	Major	Major	Major	Major
Physical_Defects_x	Moderate			
Contributing_Cause_x		Major	Major	Major
Type_of_Vehicle_x	Major			
Vehicle_Maneuver_x				Moderate
point_impact_x		Major	Major	
nWork_Area_x	Minor			
Private_vehicle_use_x		Minor	Minor	
Speed_limit_x	Major	Major	Major	
ADT_PER_LANE	Major			Major
nAVGTFACT	Minor			
Concrete_Surface	Minor			
nRural_Urban	Moderate	Major	Major	Moderate
nType_of_Shoulder				Minor
Roadway_Curve				Moderate
Sidewalk_width_group	Minor			
LIGHTCDE	Moderate			Minor
LIGHTING	Minor			
Type_Friction_Course	Minor			
Skid_Resistance	Moderate	Minor	Minor	
AUX_Lane_Num			Moderate	
Day_of_Week	Moderate			

Note: Major ($p<0.001$), Moderate ($0.001 \le p<0.01$) , Minor ($0.01 \le p \le 0.05$) and Marginal ($p>0.05$)

In the crash type models, the driver characteristics variables dominated the relative significance (see Table 5-59). In the angle, left turn and fixed object crash models driver behavior variables also had major relative significance. The ejected event was the most significant of the crash variables for fixed object, angle and left turn crashes. It was not significant for the rear-end crashes, while type of vehicle had major significance only for rear-end crashes. This suggests that vehicle protection capabilities are more important in rear-end crashes. In regards to road-related variables, speed limit and *adt* per lane were the only variables with major relative significance. The land use had moderate relative significance in the rear-end and fixed object crashes, while it was major for the angle and left turn crashes. Other road-related variables had minor or moderate relative significance.

5.4.4 Comparison of Models

The major relative significance of road-related variables in the road entity models was surprising given the weakness of the road-related variables in the exploratory analysis. The stratified sampling technique for the final analysis and the addition of road characteristics variables are likely contributors to this improvement in the injury severity analysis. The relative significance of the variables for these models suggests the importance of access management techniques, basic road, roadside and pavement design parameters. Meanwhile, the crash type models had a significant weakening of road-related variables. These are most useful to determine indirect effects of road features on injury severity (such as unsignalized intersections in tangent sections for fixed object crashes and median type for left turn midblock crashes). Also these collision types' models include a few driver, vehicle and crash effects not captured by the most general models. In general, all the models agreed on the signs of the coefficients and were this was not the case, there was a good empirical or intuitive reasoning behind these changes.

Even with the presence of a plethora of driver-, vehicle- and crash-related variables in the models, an impressive number of roadway-related factors entered the driver injury models and remained significant. In most cases, the addition of certain roadway-related variables apart from the stepwise model (such as traffic control and an interaction between skid resistance and land use) significantly improved the performance of the models.

Table 5-60: Goodness of fit Measures for the Best Candidates of the Road Entity Models

GOF Parameter	OVERALL	SEGMENT
Number of Variables	38	32
Degrees of freedom	68	60
Marginally significant coefficients	5	10
Non-significant coefficients	7	7
Sample size	107449	69887
Response severe/non-severe ratio	6.09%	6.74%
AIC	41752.15	28556.32
Hosmer-Lemeshow p-value	0.7626	0.9695
c value (area under ROC curve)	0.774	0.788
Percent Concordant	76.8	78.3
Adjusted R-squared	0.1883	0.213

In terms of statistical significance, almost all of the models had acceptable goodness of fit measures. In addition to the goodness of fit of the models, the model reliability is dependent on the goodness of fit, robustness of risk factor coefficients, and the consistency of the coefficients with previous empirical results and scientific principles. The multiple interactions of factors with intersections and the land use underscore the importance of these characteristics related to road design in the analysis of the multilane arterial safety performance. The goodness of fit measures for the road entity models (see Table 5-60) suggest a very good fit of the overall model and the segment models with 76.8 and 78.3 percent concordant, respectively. Due to the weaker intersection models, no combination of segment and intersection models could match the goodness of fit of the overall model. Considering the interpretative value of the models and the robustness and consistency of the coefficients, the overall model was the one that best captured the effects affecting the driver injury severity in crashes on high-speed multilane arterials. This analysis suggests that the overall (joint analysis) is the most reliable of the road entity models.

Table 5-61: Goodness of fit Measures for the Best Candidates of the Crash Type Models

CRASH TYPE MODEL	REAR-END SEGMENT	REAR-END PURE SEG	ANGLE UNSIG	LEFT TURN UNSIG	FIXED OBJECT SEGMENT
Number of Variables	20	16	12	14	14
Degrees of freedom	34	27	18	25	23
Marginally significant coefficients	4	3	0	4	3
Non-significant coefficients	7	6	2	4	4
Sample size	20281	14580	6273	6289	4298
Response severe/non-severe ratio	3.54%	3.48%	9.58%	9.59%	14.61%
AIC	5500.22	3891.13	3434.96	3433.34	3042.54
Hosmer-Lemeshow p-value	0.0995	0.3388	0.2267	0.7437	0.7322
c value (area under ROC curve)	0.762	0.767	0.75	0.759	0.767
Percent Concordant	75.3	75.8	74.6	75.5	76.5
Adjusted R-squared	0.142	0.146	0.184	0.192	0.224

The situation is more complex for the crash type models. Rear-end and fixed object crashes were expected to perform better in the pure segment models, while the intersection models were expected to be the best for angle and left turn crashes. The goodness of fit measures for the crash type models (see Table 5-61) showed some unexpected results. In the case of the rear-end crashes, the pure segment model performed better (adjusted R-squared=0.146) but limited the analysis by excluding crashes at unsignalized intersections which had a similar performance (adjusted R-squared=0.142). This limiting effect was seen in the coefficient robustness and thus both models are considered best candidates. The angle and left turn crash models suffered similar limitations, where the unsignalized intersection models performed better with higher adjusted R-squared values (0.184 and 0.192) than the rear-end models. In both cases (left turn and angle crashes) the pure segment models performed better than the signalized intersection models. This suggests the importance of driveway (access points) left turn and angle crashes while it underscores the weakness of the signalized intersection models. Finally, for the fixed object crashes the segment model proved to be the best candidate without any major reservations due to their favorable goodness of fit statistics and its coefficient robustness. In fact the fixed object segment model performed better than any of the other crash type models (adjusted R-squared 0.224). This validates the research methodology of separating fixed object crashes for further analysis of driver injury severity for crashes on high-speed multilane roads.

In general, the crash type analysis suggests a good fit of the models. The rear-end crashes had the weakest goodness of fit while the fixed object crashes had the best performance in the injury severity analysis. The left and angle crashes had similar performances, as expected. It is likely that better performance can be achieved if additional intersections information becomes available. The selection of best candidates does not render the other models useless. Additional information can be extracted and compared with future research from the 20 models developed for the crash types. The conclusions of this investigation were focused on the analysis of the best candidates. Considering the interpretative value of the models and the robustness and consistency of the coefficients, this investigation determined which candidates capture the most effects affecting the driver injury severity in crashes on high-speed multilane arterials. It was observed that the segment models for the rear-end and fixed object crashes as well as the unsignalized models for the angle and left turn crashes were the best performers (see Table 5-62, page 245).

Table 5-62: Summary of Best Candidates of the Crash Type Models

Entities/ Collision type	Signalized	Unsignalized	Segments plus Unsignalized	Purely segment	Overall Model
Rear-end			X	X	
Angle		X			
Left Turn		X			
Fixed Object			X		

The best candidates presented in this section might have been affected by a few factors not controlled in this investigation that should be pointed out. First, the weakness of the signalized intersection models due to the lack of intersection, signal timing/phasing and traffic data. Secondly, underlying differences correlated to land use were found that could cause some misspecification symptoms in these injury severity models. The implications of this situation are discussed in Section 5.6. Lastly, correlations among crashes and between closely spaced road features (especially intersections) were not accounted for by the statistical methods presented here. The vast majority of statistical information and conclusions presented are likely to hold true in future research judging from their agreement with past studies. However, additional investigation is also likely to uncover additional relationships

This section compared the models by their statistical goodness of fit robustness and validity with previous research and scientific principles. In addition, it suggests a course of action for future research, as more information becomes available and additional statistical methods are brought to bear on the analysis of crashes on high-speed multilane roads. By discussing some of the issues found during the course of this investigation, additional understanding of future avenues of research was achieved.

5.5 Model Considerations

Stepwise selection of main effects and interactions using statistical significance can provide a valuable contribution to model identification, especially when there are many scientifically possible interaction effects (Hosmer and Lemeshow, 2000). The stepwise method of forward and backward elimination of variables in the models, although very useful and robust, is far from perfect and requires additional analysis to properly develop the models following certain guidelines. The stepwise method in SAS uses the score test for statistical significance of the models. In addition, the AIC measure (in lieu of the likelihood ratio measure) was used to determine if the variables added to the model significantly improved the model utility.

A larger sample provides more opportunity for additional covariates to enter the model. However, these are not necessarily significant; thus careful process of cleaning out the problematic variables with large standard errors avoid problems from sparse cells, complete separation or collinearity. These issues were addressed after by analyzing the results of the stepwise regression outputs. In the case of complete (or quasi complete separation) SAS allow the iterations to continue and gives a warning, leaving the decisions of allowing variables in the model to the user. In the case of collinearity, the detection of large standard errors across one or more models justified the removal of some main effects or interactions (i.e. driver age and at-fault) that showed empirical and intuitive evidence of collinearity. The elimination of such interactions was verified with the change in Likelihood Ratio Test (hierarchical analysis) of the previous model and the modified model (without the offending interactions). In addition, a significant change in the AIC value (more than 10, as discussed previously) was verified.

The selection of interactions in the model followed the hierarchically well-formulated (HWF) model principle. The HWF model includes all interaction terms as main effects in the model, regardless of their coefficient significance, and is used in most interaction analysis applications (Jaccard and Dodge, 2004). Only two-way interactions were considered in the model, as the interpretative ability for more than two-way interactions is too complex. HWF models avoid omitted variable bias by including all the interacting variables and its main effects.

On the other hand, a few main effects with large (not unreasonably) standard errors were left in some models due to their importance as contributing factors or because the offending coefficient corresponds to the *other* category. Some of these problems are usually numerical in nature and not necessarily indicate a problem with the data. In a few cases, such as the traffic control variable of the overall model, important main effects dropped from the stepwise model after introducing additional variables. Since the 0.02 entry and stay variable requirement was introduced due to the size of the sample, the stay requirement was relaxed to 0.05 in this case to test whether this variable had a confounding effect or it was indeed a main effect. The resulting model did not introduce additional variables and the maximum change in odds ratios was 1.76%. Therefore, the traffic control variable was added to the model even though the standard error for one of its levels was large. In such large models of a joint (grouped data) analysis is practically impossible to avoid a small degree of numerical problems without dropping important variables from a logistic model. A great degree of care was taken in the data preparation and sampling so that no additional data problems would compound the numerical challenges posed when developing these models.

No major collinearity effects are suspected among the variables selected. In case of variables with unexpected coefficients, their distributions were analyzed to find whether the coefficient values were affected by collinearity or confounding effects. Even if certain degree of collinearity exists, variables should not be removed from the models to avoid omitted variable bias. Meanwhile, combining different variables may cast doubts on the data preparation process or the theorized variable effects or processes (Menard, 2001).

One strategy to evaluating the importance of predictors in logistic regression is to evaluate their odds ratios. Those predictors with the larger changes in outcome are considered the most important (Tabachnick and Fidell, 2001). This approach was followed in the discussion of the models, by comparing the magnitude of the effects. Another approach to evaluate the importance of predictors is the Type III tests in the SAS output, which are explained in Section 5.4.3.

Some important variables, such as land use became non-significant in presence of many other roadway-related variables and the interaction with skid resistance. These results suggest that the land use is represented by other road characteristics including access management. The land use was significant in every road entity model except in the pure segment. The strong correlation between the skid resistance numbers and the land use may be related to the operating speeds and design differences between urban and rural areas. The nature of these relationships deserves special attention and should be investigated further in the future.

5.6 Alternative Corridor Injury Severity Model by Land Use

5.6.1 Misspecification Analysis

There was some evidence of a misspecification problem with respect to the most important exposure variable (*adt* per lane) in crash analysis for the signalized intersection model. Such a development questions the validity of performing traditional road entity analysis for arterial corridors where spatial (and possibly temporal) correlations exists among the different road entities. The overall model for the arterial corridor provided the only choice that would include the exposure variable for all the crashes in the corridor. This situation was not totally unexpected since the exploratory models for signalized intersections did not perform well when the single and multiple crashes were combined, among other issues found during the exploratory analysis. It is worth nothing that in the models in previous studies (Wang et al., 2006; Abdel-Aty and Wang, 2006) the *adt* per Lane variable included traffic volumes from both the major and minor roads. The lack of minor road traffic volume data in the present investigation is not considered a factor to the misspecification because the *adt* per lane variable considers only the volume and lanes of the major road, thus the ratio is valid for the major component of the intersection. The coefficients values might be biased, but still efficient. In a study of crash severity levels at signalized intersections, Abdel-Aty and Keller (2005) found that the major road *adt* was significant in the model, as will be described.

Another possible cause for the misspecification may be the lesser robustness of the logistic model vs. the family of generalized linear models. However, the results in the literature instead point to a possible high correlation between the land use and *adt* leading to possible model misspecifications for spatially correlated intersections. Obeng (2007) successfully entered the *adt* per lane variable in a binomial logit model for 303 signalized intersections in the city of Greensboro NC. The injury severity analysis was successful when the log of the *adt* of the intersecting roads was entered into the model. All of the intersections were in an urban area. On the other hand, Krull (2000) used logit models to analyze injury severity of single vehicle crashes in rural and urban areas. The *adt* variable was excluded due to a high correlation (0.533) with the rural functional class. Finally, in a study of crash severity levels at signalized intersections, Abdel-Aty and Keller (2005) achieved gains in variable information in the intersection characteristics with a complete dataset model with major road number of lanes, left and right turn lanes, division on minor road, and *adt* on major road. In this study, ordered probit models similar in robustness to spatial correlation to the logistic models were used in the analysis. Data for the minor road (*adt* and number of lanes) were available and tested, but was not significant in the model. Multiple years (2000-2001) were used and different jurisdictions in rural and urban areas were selected with intersections of all speed limits and number of lanes.

Various issues were found during the exploratory and final analyses that were related to the land use, including earlier issues with the skid resistance variable. Therefore, a final test of the combined (all) intersections and signalized intersection models was developed. Four models (two rural and two urban) were fit including the log_ADT_per_lane variable. The results indicated that for all intersections, the *adt* variable was significant for the rural area (p-value=0.0454), while it was highly insignificant (p-value=0.6046) for the urban area. Meanwhile, for the signalized intersections, the *adt* per lane variable was marginally significant (p-value=0.1756) for the urban area and highly insignificant (p-value=0.7115) for the rural area. These results show enough evidence of misspecification for the intersection models. The resampling including involvements from drivers 1 and 2 and the careful selection without substitution to avoid crash and road data repetition was successful. It has allowed discovery of the root causes of various issues found since the exploratory analysis. In addition, the correlation problems of some variables did not degraded the signalized intersection model performance due to its increased robustness. The implications for both the modeling strategy and the validity of the model results are discussed in Chapter 6.

5.6.2 Corridor Injury Severity Model by Land Use

The original research strategy did not include analysis of driver injury severity by land use. However, after examining the results of the different models, it was clear that alternative models analyzing all crashes by land use should be investigated. The goodness of fit of the models improved significantly (see Table 5-63, page 252). The AIC values improved dramatically, even better than the other two combinations by road entity models (all intersections and segment plus unsignalized) even with a larger sample and number of variables. In addition, the severity proportions have a larger difference than across the road entity models, which denotes a different injury severity distribution. It is clear from these data that the proposal for corridor level injury severity analysis should contemplate land use as unit of analysis among other possible options. In this investigation, it showed to be superior to the traditional road entity options. It also shows the model limitations, especially in the urban area crashes. In this section, the major differences between the three models will be briefly discussed. Details about the model coefficients and standard errors can be found in Appendix F.

Table 5-63: Goodness of fit Measures for the Final Analysis Injury Severity Models by Land Use

GOF Parameter	OVERALL	ALL_RUR	ALL_URB
Number of Variables	38	31	33
Degrees of freedom	68	57	61
Marginally significant coefficients	5	8	6
Non-significant coefficients	7	8	8
Sample size	107449	50953	56496
Response severe/non-severe ratio	6.09%	7.44%	4.88%
AIC	41752.15	22609.71	19073.417
Hosmer-Lemeshow p-value	0.7626	0.5906	0.3801
c value (area under ROC curve)	0.774	0.780	0.763
Percent Concordant	76.8	77.5	75.5
Adjusted R-squared	0.1883	0.2055	0.1646

With regards to the driver-related variables, the injury severity model by land use introduces some very useful information about driver characteristics (see Table 5-64, page 254). Driver age had the greatest variation for the negative effect on very old drivers in rural vs. urban areas (OR=1.822 vs. 1.394), which should have important consequences in safety guidelines for rural areas. Most safety literature regarding older drivers focused on intersections and lane guidance in urban areas. Gender (female drivers) denotes negative effects (OR=1.256) in rural areas, while in urban areas it is not significant. A similar situation happens with the interaction between female drivers and seat belt usage, which has a negative effect (OR=1.257) compared to male drivers using seat belts only in the rural area. The speeding and contributing cause (aggressive driving and driving under the influence) show increased negative effects in the rural area compared to the urban areas. The at-fault driver effect remains almost equal, it would be worthy of additional investigation how is it that driver behaviors have differences in their effects by land use yet the at fault driver effect remains similar in urban and rural areas. The residence code is only significant in urban areas, which is logical since non-residents are less likely to travel in rural areas without touristic attractions. Physical defects do not follow the trend of the rest of the variables. It seems that drivers in urban areas have a disadvantage if they loose control compared to in rural areas, which might be due to the additional traffic (which makes a out of control crash more likely) and limited roadside clear zone in urban areas.

In regards to crash types, the only negative effect that increases in rural areas is for the head-on collisions. The angle crashes have similar odds ratios, while for left turn crashes drivers traveling in urban areas tend to have higher severe injury odds ratio when compared to the rural areas. For sideswipe crashes, there is a positive benefit (OR=0.683) for crashes in rural areas and no significant effect in urban areas. Lastly, there is a larger negative effect for driver severe injury in fixed object crashes in urban areas (OR=2.374) when compared to rural areas (OR=1.531). Other crash variables have an increased effect in urban areas as well, such as off roadway (recall the fixed object crash at unsignalized intersections model), multivehicle crashes and multivehicle crashes at or near intersections. For the type of vehicle, the LTVs similar odds ratios in both land uses. Meanwhile, heavy trucks and especially motorcycles and bicycles have a negative effect in urban areas.

Table 5-64: Odds Ratios for Driver-, Vehicle- and Crash-related Variables in Final Analysis Models by Land Use

Variable	Level	Rural	Urban	Overall
Driver age 80-98 (vs. 25-64)	**5**	1.822	1.394	1.621
Driver age 65-79 (vs. 25-64)	**4**	1.412	1.437	1.422
Driver age 20-24 (vs. 25-64)	**3**	0.792	0.755	0.779
Driver age 15-19 (vs. 25-64)	**2**	0.761	0.781	0.767
Speeding (Unknown vs. not)	**2**	0.747*	1.037†	0.863†
Speeding (Yes vs. Not)	**1**	0.450	0.335	0.409
Gender (Female vs. Male)		1.256	1.190*	1.217
Seat Belt Used (vs. no)	**1**	0.265	0.358	0.303
Gender*Seat Belt used	**1**	1.257	1.216*	1.245
Other Cont. Cause (vs. none)	**4**	1.641	1.544	1.605
Aggressive Driving (vs. no improper)	**3**	1.796	1.655	1.748
Alcohol/Drug use (vs. no improper act)	**2**	1.609	1.588	1.593
At Fault driver (vs. not)		0.541	0.523	0.538
FL Resident (vs. not)			1.307	1.175
Physical Defects (vs. not)		1.454	1.643	1.535
Driver Ejected (Yes/Partial vs. No)		4.591	3.928	4.270
Other collision type (vs. rear-end)	**7**	1.118*	1.064†	1.097
Fixed Object (vs. rear-end)	**6**	1.531	2.374	1.810
Sideswipe (vs. rear-end)	**5**	0.683	0.883†	0.779
Left Turn (vs. rear-end)	**4**	2.034	2.540	2.242
Angle (vs. rear-end)	**3**	1.789	1.801	1.784
Head-On (vs. rear-end)	**2**	3.115	2.755	2.875
Other vehicle type (vs. automobile)	**5**	0.743*	0.794†	0.756
Bike/motorcycle (vs. automobile)	**4**	0.773	1.524	1.050†
Trucks/buses (vs. automobile)	**3**	0.306	0.455	0.357
Van/Light Truck/Pick up (vs. automobile)	**2**	0.823	0.821	0.820
Point of Impact (driver side vs. not)		1.035†	1.170*	1.091†
Point of impact*Speeding	**2**	1.247†	1.209†	1.240*
Point of impact*Speeding	**1**	1.441	1.390	1.412
Off Roadway (vs. not)		0.590	0.670	0.613
Off Roadw*Speeding	**2**	0.872†	0.681	0.783*
Off Roadw*Speeding	**1**	1.222*	1.513	1.289
Off Roadw*Multivehicle	**1**	2.046	1.882	2.033
Work Area (Entered vs. none)	**3**	0.747	0.997†	0.826*
Work Area (Nearby vs. none)	**2**	0.741	0.772	0.750
Multivehicle	**1**	0.475	0.519	0.469
Intersection*Multivehicle	**1**	1.214	1.737	1.476

Notes: * Effect is marginally significant ($p<0.20$); † Effect is not significant ($p\geq0.20$)

In regards to the roadway-related variables, the major changes were noticed in the traffic control variable (stop vs. none), which is marginally significant for the rural area only (see Table 5-65, page 257). The perceived benefits of access management in crash severity are not significant for the rural areas, even when testing the other class levels (2-4). Only class 5 is marginally significant. Additional research is required to better determine the relationships between access management and crash severity in rural corridors. One feature not available in the road inventory database is the driveway density, which becomes is presumably an unobserved factor. The type of shoulder suggests an advantage of unpaved shoulder over paved shoulders in rural areas and a disadvantage of curb shoulder over paved shoulders in urban areas. It seems that the unpaved shoulder in rural areas may be correlated to better clear zone, but additional research is required.

In rural areas, conditions on roads with lanes between 10 and less than 11 ft wide seem to be more favorable in terms of injury severity than widths less than 10 ft or between 11 and 12 ft. Meanwhile in urban areas, sections lanes more than 12 ft wide have the most favorable conditions, according to this model. In rural areas narrow lanes would likely reduce operating speed (a major cause of severe crashes in rural areas) while in urban areas, increased heavy vehicle traffic would likely benefit from wider lanes. This information is much better than that of the overall model. Roadway curves are a great concern in urban areas, contrary to the traditional approach in two lane roads due to the fixed object crashes in unsignalized intersections on tangent road sections. On the other hand, the land use models suggest that wider sidewalks are more likely to benefit drivers in rural areas, while the effect is not significant in urban areas. This might partially be due to other road conditions (improvements) in sections were sidewalks are built. Even the sidewalks and changed landscape might act as a traffic calming effect in these rural sections of multilane arterial corridors.

The high-mast (LIGHTING) effect is only sign cant in urban areas. These high masts are usually placed at or near interchanges and the roadside conditions in urban areas vary in such a way that these poles may become a hazardous fixed object even when protected by guardrails. High masts are no longer used in many places; more information about the locations of these high-masts is needed to determine if these are already hazardous locations. However, these is a large benefit to have full lighting (OR=1.342) when compared to partial lighting (OR=3.280). For the conventional light poles, the odds ratios suggest that partial lighting in rural areas has a significant benefit, while full lighting in urban areas have a disadvantage. Additional research is needed to determine whether the closely spaced poles are the culprit in urban areas or if other effects (roadway curves with increased lighting or hazardous locations where additional luminaries are installed).The type of friction course is only significant in the urban area and type course 5 becomes insignificant. More information is needed about the proportions of these older friction courses (FC-2 is a dense graded, no longer effective) in the roadway inventory data. The relationships of skid resistance (as discussed previously) and day of week (weekend vs. weekday) are similar for both land uses.

The driver injury severity analysis by land use presented here is proposed as a framework for future research involving severity analysis in high-speed multilane arterial corridors. Additional road and land use information would be needed, as well as more powerful analysis methods to fully exploit the potential of this manner of corridor level analysis. A modeling scheme by road entity and crash types is recommended to test the reliability of injury severity models by land use.

Table 5-65: Odds Ratios for Roadway-related Variables in Final Analysis Models by Land Use

Variable	Level	Rural	Urban	Overall
Speed limit (40-45 vs. other)	**1**	0.663	0.701	0.676
ADT per Lane (thousands)		0.971	0.976	0.972
Avg Truck Factor (percent)		1.010	1.013	1.011
Traffic Signal (vs. other control)	**3**	1.127*		1.128
Stop/Flashing (vs. other control)	**2**	0.968†		0.997†
Non applicable (vs. class 2,3,4)	**9**	1.091*	0.991†	1.030†
Access class 7 (vs. class 2,3,4)	**7**	1.116†	0.657	0.781
Access class 6 (vs. class 2,3,4)	**6**	0.965†	0.785	0.833
Access class 5 (vs. class 2,3,4)	**5**	0.914*	0.868	0.879
Urban area (vs. Rural)		N/A	N/A	0.879
Curb Shoulder (vs. Paved)	**3**	0.980†	1.252	1.089
Unpaved Shoulder (vs. Paved)	**2**	0.887	1.124*	0.967†
Lane width (vs. 11 ft ≤ width ≤ 12 ft)	**Desc.**			
Lane width <10 ft	**4**	0.886*	0.807	0.827
10 ft ≤ lane width < 11 ft	**3**	0.673	0.865*	0.815
Lane width > 12 ft	**2**	0.804	0.854	0.810
Roadway Curve (vs. non curve)		1.196	1.536	1.306
Sidewalk width ≥ 6 ft (vs. < 4 ft)	**3**	0.739		0.791
4 ft ≤ Sidewalk < 6 ft (vs. < 4 ft)	**2**	0.797		0.851
High Masts (full vs. none)	**Y**		1.342	1.331
High Masts (partial vs. none)	**P**		3.280	3.506
Full Non-High Mast (vs. none)	**Y**	0.899†	1.304	1.129
Partial Non-High Mast (vs. none)	**P**	0.727	0.948†	0.821
FC-3, FC-6, N/A (vs. FC-2)	**9**		0.803	0.975†
Friction Course 5 (vs. FC-2)	**5**		0.805*	0.831
Friction Course 4 (vs. FC-2)	**4**		0.784	0.918
Friction Course 1 (vs. FC-2)	**1**		0.621	0.736
Intersection			0.630	0.831*
Intersection*Urban Area		N/A	N/A	0.862
Skid Resistance (1≤FN<35 vs. FN≥35)		1.178	1.126	1.198
Urban Area*Skid Resistance		N/A	N/A	0.919*
Weekend (vs. Weekday)		0.902	0.909	0.906

Notes: * Effect is marginally significant ($p<0.20$); † Effect is not significant ($p \geq 0.20$)

CHAPTER 6. CONCLUSIONS

In the preliminary analysis, the differences between intersection and non-intersection involvements were explored with cross tabulation tables. In the exploratory analysis, the involvements in high-speed multilane roads were analyzed using logistic regression models by road entities. A massive data preparation effort was undertaken to correct deficiencies found during the exploratory analysis. Preliminary analysis of the involvements from vehicle sections 1 and 2 showed evidence of possible driver selection bias. To avoid bias, it was decided to choose a random sample of multivehicle crashes using the first two driver sections without substitution. Single vehicle crash driver involvements were added to this sample without repeating crash records. The final analysis consisted of six road entity models and twenty crash type models. Both the data preparation and sampling were successful in allowing a robust dataset. The overall model was the best candidate for the analysis of driver injury severity on high-speed multilane roads. Driver injury severity resulting from angle and left turn crashes were best modeled by separate unsignalized intersection crash analysis. Injury severity from rear-end and fixed object crashes was best modeled by combined analysis of pure segment and unsignalized intersection crashes.

The most important contributing factors found in the overall analysis included driver-related variables such as age, gender, seat belt use, at-fault driver, physical defects and speeding. Crash and vehicle-related contributing factors included driver ejection, collision type (harmful event), contributing cause, type of vehicle and off roadway crash. Multivehicle crashes and interactions with intersection and off road crashes were also significant. The most significant roadway-related variables included speed limit, *adt* per lane, access class, lane width, roadway curve, sidewalk width, non-high mast lighting density, type of friction course and skid resistance. During model building some misspecification symptoms appeared due to major differences in road and crash types by land use. Two alternative models of crashes in urban and rural areas were successfully developed. The models by land use were substantially better than any other combination by road entity or the overall model, as indicated by their AIC values. Their coefficients were substantially robust and their values agreed with scientific or empirical principles. Injury severity models by land use should be further investigated following the road entity and crash type modeling scheme used in the present investigation. A framework for injury severity analysis and safety improvement guidelines based on these results is presented.

6.1 Analysis Methodology Implications

The main goal of this investigation was to find an appropriate method for driver injury severity analysis of crashes occurring on high-speed multilane arterials. There are broad implications for the evaluation of the safety performance of arterial corridors. Disaggregated analysis by road entity and crash types allowed a complete picture of the contributing factors affecting driver injury severity and comparison between different models. From this analysis, we can assess the model reliability and present recommendations for future research.

First, the road entity models that had the best goodness of fit measures were observed for the overall, segment and pure segment models. These three models exhibited the best calibration, percent concordant and adjusted R-squared values. The combined models (segment and overall) provided the best set of contributing factors. In terms of coefficient robustness, the overall model had at least 84% of its variables with major or moderate relative importance. The number of non-significant coefficients also compared favorably to the other road entity models. Additional intersection data (movement counts, signal timing, geometry) might improve the models related to intersections. However, the overall model included the variables related to intersections and interactions with intersection presence, which represented most intersection effects, with the notable exception of red light running.

Secondly, the models by crash type had more complex relationships. For the rear-end crashes, the segment and pure segment models had a better fit than the other models. The segment model included all variables present in the pure segment model and the overall model except a few variables with minor or moderate relative importance. The pure segment model on the other hand did not have the land use variable, which has major significance. Modeling of rear-end crashes with a segment model is expected as most segment crashes are rear-end. The major disadvantages were the absence of the lane width (which has moderate relative significance) variable and the low calibration power as shown by the Hosmer Lemeshow p-value. For the angle crash models, the unsignalized intersection model had the best fit due to the very low calibration power of the segment and overall models. The major disadvantage of the unsignalized intersection model for angle crashes was the absence of the access class and vision obstruction variables, which were significant for signalized intersections and pure segments, respectively. For modeling driver injury severity in left turn crashes, the unsignalized intersection model was the best fit. However, in this case all models had excellent calibration. The major disadvantage of this model was the absence of the access class variable (moderate relative significance). For fixed object crashes, the segment model was the best fit without major disadvantages. The combination of pure segment and unsignalized intersection fixed object crashes resulted in excellent calibration.

In conclusion, the overall model exhibited better model reliability when considering goodness of fit and coefficient robustness. For rear-end crashes, the segment model was more reliable, with a few reservations. In the case of angle crashes and left turn crashes, the unsignalized intersection models demonstrated better performance, with the exception of the access class variable, which was significant for left turn and angle crashes in signalized intersections. For the fixed object crashes, the best injury severity analysis modeling resulted from the segment model without reservations. In general, models with higher proportions of severe crashes had a better reliability. The overall model captures almost all of the significant effects, while the angle, left turn and fixed object crash models capture a few specific effects (i.e. fixed object crashes in unsignalized intersections on road curves) important to the safety performance of high-speed multilane arterials.

6.2 Significant Factors Conclusions

This severity analysis has shown many characteristics that distinguish the contributing factors to driver injury severity in high-speed multilane arterials to those for crashes on other road types. Next, some additional implications of some of the findings are discussed. A group of findings are only preliminary, as they require additional research to prove the nature of these relationships. Most of the discussion will focus on roadway-related features and their effects on driver injury severity.

In locations with full lighting (higher density) there was a negative effect on driver injury severity. This is probably affected by the selectivity of higher densities of poles near intersections. Further research is needed to prove whether clustering among hazardous locations or whether there is a more systematic symptom. Previous research discussed previously identified driver's overconfident behavior offsetting safety benefits of certain design features. The key advantages of the overall model to analyze the driver injury severity resulting from crash involvements on high-speed multilane arterial corridors are realized in the roadway-related variables.

The off-road crash interactions emphasize the role of speeding as a safety hazard in high-speed multilane roads and points to a possible relation between driveway crashes (multivehicle and off road) and increased injury severity. The current literature on the relationships between driveways and crash severity on multilane roads is limited. Hauer (2004) found that commercial driveways contributed to increased injury in on the road crashes in four lane undivided (including TWLT) urban segments. No significant relation between driveways and off-road crashes was found in those segments. A recent study by Lui et al. (2008) showed that the presence of U-turn bays in between signalized intersections and increased separation distances between driveways and U-turn locations in divided urban and suburban roads reduced total and angle crashes. However, no literature was found that addressed the relation between driveways (or unsignalized intersections) and safety performance of rural multilane segments. The additional findings about the interactions of road safety characteristics (such as skid resistance) and rural areas indicate that high-speed multilane roads in rural areas should be further investigated separated from the urban sections due to their major differences demonstrated in this investigation.

The lowest odds ratios were for the lanes between 10 and less than11 ft in rural areas and more than 12 ft in urban areas. This difference might be attributed to the wide lanes for bicyclists in urban areas. A recent observational study by Hunter and Feaganes (2004) recommended

FDOT to convert 14 ft wide curb lane conversions to bicycle lanes 3 ft wide and 11 ft vehicular lanes. The sites selected for the study were in multilane facilities with curb and gutter (except for one site) and speed limits between 40 and 45 mph. Their conclusions were based on rates of vehicle encroachment on the bicyclists shared path vs. the designated 3 ft lane. No additional literature which included crash analysis of motor vehicle traffic was readily available. This and other bicycling countermeasures may have adverse effects on severity of crashes on multilane arterial corridors, as suggested by the present investigation. It is imperative that additional crash analysis is performed in these sites to evaluate the safety effects of these lane conversions.

The shoulder width and median width did not enter the overall models. Even in the rural area model, only shoulder and median widths were marginally significant, while in the urban area only median type (raised vs. paved) was only marginally significant. Also, the results from this investigation also placed an important role on median design and unsignalized intersection characteristics. These results agree with the conclusions of Gattis et al. (2005) and Eisele and Frawley (2005), as discussed previously. Traditionally auxiliary left and right turn lanes on the major road have been found to provide increased safety benefits. However, in multilane arterial corridors additional conflicts with through movement seem to have the most significant effect on severity. Among those are rear-end crashes on a median opening (encroachment of through lane). This suggests that access management and intersection design have a more important effect on high-speed multilane arterials than other road elements that have traditionally been targeted for safety improvements. The safety performance of high-speed multilane arterials is dependent in a greater degree on traffic (turn movements and crossings) control and access point spacing to reduce weaving.

The presence of multiple speeding interactions with other crash-related variables emphasizes the role of speeding as a safety hazard in high-speed multilane roads. These types of roads have a degree of complexity, lack of uniformity (from one land use to the next) and hazardous crossing points that give much less room for error or unsafe behavior for drivers, when compared with limited access roads. Yet, due to their nature and enforcement funding structure (NHSTA funding) which favors freeways, traffic enforcement is notably less than in other types of roads. There are signs of recent changes in the allocations for traffic enforcement.

Number of approaches at intersections was tested in the exploratory and final models, but was not found significant in any model. The variable did not enter the final models, a significant finding. The implications that the numbers of approaches at high-speed multilane road intersections are not significant in regards to the driver injury severity include design considerations. First, the intersections in these types of roads with more than four approaches were rare. Second, other variables, such as major road lane width and *adt* per lane, which relate to the intersection size and operation, were found significant in the models. This suggests that certain intersection design features act as significant factors of driver injury severity, while the more general characteristic of number of approaches, which may increase the number of crashes, does not have a significant effect. Recall that unsignalized intersections, which can be considered as three leg in a divided road, were found to have negative effects due to other design considerations, such as roadway curve and major road lane width (for left turns).

One of the most important road design parameters is the speed. Speed limits generally reflect a decrease of 5 to 10 mph below the speed limit. However, there are many changes on the roads that degrade the design speed. The development of roads in rural areas does not necessarily follow the pace of business development along the arterial corridors. Access management is currently included in the retrofit plans for high-speed multilane arterials.

An interaction variable of speeding and point of impact was tested in the models to prove whether this drivers speeding would be usually at an advantage when hitting other vehicles. Since the speeding indicator is computed using the estimated speed reported by the police officer only for certain types of crashes, there are more missing data than for any other variable. It is more likely that the police officer reports an estimated speed for a severe crash requiring a thorough investigation. Thus, it was deemed pertinent to include this variable (with one level labeled unknown) for its perceived significance in severe crash outcomes. The point of impact and speeding interaction was significant in both the rural and urban models. There was a negative effect, likely due to angle crashes with impacts on the speeding driver's side. Speeding also played an interaction role with off roadway crashes, which tend to be severe.

There is a tendency of increased total and severe injury involvements at locations with older friction courses, as shown in both Figure 5-6, page 186, and Figure 5-7, page 187. Decreasing skid resistances of older friction courses (polishing effects) is an important concern for skid hazard prevention programs. However, roads under wet pavement hazards are considered when at least 25% of the crashes are related to wet pavement. If there is a systematic decrease in friction resistance on high-speed multilane corridors with older friction courses, it is not necessarily captured at the district level. Additional injury severity models and other systematic crash analysis should describe these tendencies in more detail.

In multilane arterial corridors, the traffic conflicts seem to have an important effect on the safety performance. In a given roadway curve, the arc length between two points is longer than its subtended length. If the access class standards do not change by horizontal degree of curvature, the density of access points (in two-dimensional space) and traffic conflicts tends to increase on curves. An example of this situation is shown in Figure 6-1. Note the closely spaced turn bays on the tangent. In addition to vehicular control, visual distraction (signs, among others) and visibility obstructions (i.e. shrubbery) increase the level of driver discomfort in urban curves. Additional traffic conflicts in closely spaced access points make the curved roadway sections in multilane roads difficult to negotiate, especially lane changes to reach left or U-turn bays from a driveway.

Figure 6-1: Curved Section of SR-423 (Lee Road) Next to the I-4 Exit in Orlando (Source: Microsoft Virtual Earth)

6.3 Analysis Success and Limitations

This study attempts to address a recent paradigm shift in arterial crash analysis: a joint analysis of different road entities with the corridor as unit of analysis. The methods tested included involvements severity analysis models of all the crashes in the corridor, by road entity and by crash types. The logistic regression analysis proved to be an effective, flexible method robust enough for the goals of this research. The results showed the potential of injury severity analysis of all crashes in arterial corridors, while they showed a flaw in the analysis method. The injury severity had a direct correlation with the land use, which was also a covariate in the model. This caused some issues with important variables which were addressed by modifying the all involvement injury severity analysis to analysis by land use. The results were successfully applied to the multilane arterial driver, vehicle and road conditions and compared with the overall model, which could not discriminate the effect of changing road conditions on driver injury severity.

An argument can be made that by analyzing the different effects of land use separately, the traditional analysis by road entity can be effectively complemented by overall crash models when analyzing multilane arterial corridors. Additional injury severity analysis of models by land use, crash types and road entity is recommended to prove whether this is true. A new proposed analysis of a set of road entities with a common land use replaces the previous notion that points of traffic conflict and road segments are to be treated as isolated entities in crash analysis. This analysis method may still be true for high-speed two-lane roads, which tend to be in rural areas with intersections not as closely spaced, but it has not been very effective in analyzing high-speed multilane arterials, as indicated in this investigation. Recent literature has focused on this area and promises to change the engineering focus in treating arterial corridors nationwide.

This analysis has some limitations that need to be accounted for in the results. This analysis only addressed driver injury severity, while most useful, does not include other effects on occupants and/or pedestrians. There is another limitation in clustering the injury severity data into two categories, which was the goal of this research. These results may be extended by accounting for the correlations among different driver injury severity levels. Also, accounting for correlations between involvements in the same crash will allow finding additional information from the interactions between vehicles, which is expected to improve the analysis of multivehicle crashes, such as angle and left turn.

A high order of multicollinearity is to be expected when a complete set of independent variables is used in a full-model regression analysis. Stepwise regression analysis does not adjust adequately for presence of collinearity (i.e., nonorthogonality) among the set of predictors and may generate inefficient or incorrectly signed parameter estimates. This limitation was handled by using stepwise regression as an exploratory tool and relying on empirical evidence to add (or remove) important variables and testing the model specification using the Hosmer-Lemeshow test. Additional variable interactions were tested in order to improve the interpretative power of the models. Most issues were dealt with in a timely way by applying the basic principles of categorical data analysis.

One recent injury severity study separated 14 types of crashes with different characteristics to alleviate and minimize biases inherent in a joint model (Ulfarsson and Mannering, 2004). In a recent study of the effects of rural highway median treatments and access (Gattis et al, 2005) the findings suggested that additional study should be performed to examine any correlations between median type and land use. Another study (Eisele, 2005) suggested additional study of median safety impacts over a broad range of geometric conditions in long corridors. The investigation presented here attempts to contribute in filling the gap in the general knowledge of the important safety effects (on driver injury severity) and suggest an effective way of capturing those effects in an arterial corridor. This study could not account for the spatial correlation among road features within an arterial corridor due to the limitations of the statistical method used for analysis. However, the amount of analysis and model specifications tested and satisfactory results for a systematic injury severity analysis of high-speed multilane arterial corridors should serve as the basis many avenues of future research.

6.4 Crash Data Limitations

The quality and stability of the injury severity analysis models are dependent on the quality of the dataset available. The data preparation efforts were successful in providing a series of models that were not possible in the exploratory analysis. Sacommano, et al (1994) and others have arrived to the same conclusions. A major limitation in all crash analyses are the missing data, which tend to be from minor crashes. There was a systematic effort in the investigation to maintain the severe crash proportions as close as possible to the original population to avoid possible bias in the models. Missing information appear to be random result of data entry error or incomplete reporting and not systematic. About 45% of the data were excluded due to missing information. Appropriate chi-square tests were computed for the excluded data in the final sample with no significant associations between excluded data and the injury severity variable.

Another possible source of error for the intersection models is the 250 ft influence area used for all intersections in Florida. There is a selectivity bias due to smaller intersections being assigned crashes occurring farther than their influence data and larger intersections without all the crashes that occurred within their influence area, as discussed in Wang et al (2008). This situation is not a concern in the final models by land use.

An additional 16% of the final sample (20,897 crashes/involvements) was lost due to a design feature of the RCI data that were merged. The roads use a linear referencing systems (LRS), which links adjacent sections with the same milepost (i.e. the end point of one section equals the begin points of the next section). At some intersections, the situation is more complicated, up to four sections might meet at one point. For large amounts of crash data there is no practical way to determine which side of the section border the crash belongs to. There was no evidence of systematic data loss. Improvements in crash location using coordinates and GIS might alleviate this problem.

6.5 Suggested Safety Guideline Framework

The result of this investigation suggests that a series of coordinated safety guidelines be proposed to the FDOT to implement in two simultaneous stages. One stage would impact road construction and retrofits. Meanwhile, the second stage will impact access management and maintenance activities on high-speed multilane arterial corridors. Today, there exist many guidelines that might conflict or not have enough information to allow the stakeholders make coordinated improvements at the state level. Once uniformity is achieved at the state level, other jurisdictions are likely to follow these guidelines. One of the main concerns that this analysis raised was the unsignalized intersections (not driveways) in urban arterials. Roadside and visibility issues have negative effects on the drivers' injury severity. Not enough current information is available in the safety literature focusing on these road features. Recent efforts to improve new intersection design are likely to have a positive impact. However, modifying existing designs system wide will likely have the greatest effect on the safety performance of the arterial corridors. A few low cost improvement examples are discussed next.

Unsignalized intersections on high-speed multilane corridors are sometimes difficult to identify when traveling on the major road. On the other hand, heavy traffic volumes make these intersections difficult to cross, even with the presence of a median. Some improvements in signage letter size and lane assignment signing at large intersections have been slower in state roads compared to county roads in the Central Florida area. In some other parts of the state (especially South Florida) this situation is a little different. Low-cost countermeasures, such as signing, when applied uniformly have a high impact reducing traffic flow disruptions and improving safety. Additional access management measures are needed to improve the safety performance of the high-speed multilane roads with the worst safety records (i.e. SR-50 in Central Florida). These will also improve sight distances for older drivers at intersections. In addition, clustering of closely spaced unsignalized intersections increase traffic conflicts and decrease gap availability for downstream unsignalized intersections, which is a great problem for older drivers. It seems that the permitting department plays a crucial role and there is a concern about recent access construction on state roads that do not comply with the guidelines set forth in the Median Manual.

In addition, traffic signal coordination for the corridors will not only improve travel times, but may reduce speeding and other forms of aggressive driving. If the traffic signal optimization balances well the needs of major and minor roads, driver frustration would be reduced and better gaps would be available for the unsignalized intersections in the corridor. Some of these benefits have been measured, some are difficult to measure. Better signal coordination should reduce speeding and red light running. The red light running drivers are 143% more likely to sustain sever injury, while speeding drivers are less likely (40%) to sustain severe injury when compared to non-speeders. However, these are individual factors (all else held constant) and their combination will likely yield a negative result for those drivers hit by speeders. The combination of contributing factors in the models developed in this investigation suggests that an integrated focus on the safety performance of arterial corridors yields reductions in driver injury severity. Likewise, an integrated approach to a combination of engineering countermeasures on these corridors is expected to yield the best safety improvement results.

Lane width might be increased in urban areas by reducing the shoulder width, were practical. In this investigation, shoulder width was only significant for the angle crash model. The difference in odds ratio between shoulder widths of 6 ft and 10 ft is very small compared to the system wide benefits of wider lanes in urban areas, especially those with higher truck traffic volumes. On the other hand, reducing lane widths in rural areas to 11 ft is suggested by the injury severity models to have a positive benefit. Additional research is needed to further prove these relationships before a testing program that should confirm the research findings

Drivers of vans, light trucks (LTV's) and pickups involved in crashes at high-speed multilane roads are 82% as likely as passenger car drivers to sustain severe injuries. Additional strategies to improve direction finding, lane assignment and traffic signal visibility to passenger cars driving behind LTV's may improve their drivers' safety performance. This will also help older drivers as well. Landscaping clear zone on arterial corridors will improve visibility, especially in unsignalized intersections on tangent road sections.

These examples serve to illustrate a proposed framework of coordinated efforts system wide to continuously improve the arterial corridors. Some of these examples use proven strategies to alleviate negative effects found in research. Others will require additional research and testing before implementation. Situations such as the bicycle lane assignment discussed previously must be further evaluated to make sure that an intended benefit for one group of road users does not negatively affect others. Evaluation must be timely to prevent new patterns of safety risk to rise.

6.6 Implications of Analysis Methodology and Results on Road Design and Treatments

This investigation followed the traditional injury severity analysis approach and compared its results to analysis by crash types and joint models combining crash involvements occurring at different road entities. It has been pointed out in the literature that the conditions in arterial corridors are interrelated to the land development and traffic conditions differently than for other types of roads, such as freeways. This research has demonstrated some benefits in the joint analysis of crashes at different locations, such as intersections and road segments. This does not negate the differences in crash patterns at intersections and road segments, as shown in the crash type models. Rather, it serves as proof of a relationship between the different crash locations. The exact nature of this relationship will surely be the topic of future research, but certain very useful contributing factors have surfaced in the course of this investigation can be used to point to some implications on the safety effects of design and operation parameters on arterial corridors.

Some of the most important implications of the contributing factors found in the models are related to driver, crash and vehicle factors. These include the driver age, gender, seatbelt use, ejection, speeding, drunk driving, aggressive driving and type of vehicle. Their respective safety effects (on driver injury severity) and implications are summarized in Table 6-1, page 274. One major implication of the injury severity analysis was that the joint analysis of crashes did not degrade the significance of driver, crash or vehicle factors. Another finding was the feasibility of combining different factors with roadway characteristics to improve the injury severity analysis. Some of the implications summarized point to enforcement and education strategies, as well as road design. Additional research is needed to pinpoint the nature of these strategies.

Table 6-1: Safety Effects and Implications of the Most Important Driver, Crash and Vehicle Contributing Factors

Contributing Factor	Safety Effects	Implications
Driver Age Groups	Young Drivers- positive effects	Further research in relation between driving patterns and crash involvements
Driver Age Groups	Older drivers- negative effects, larger in rural areas (land use model)	Challenging road environment- design strategies for older drivers should be continued, especially in rural areas
Driver Ejection	Large negative effects on injury severity	Emphasis on short trip seat belt use
Gender	Females have large negative effects, even with seat belts	Further research in relation of type of vehicle used, driving patterns and crash involvements
Seatbelt use	Best positive effect	Enforcement strategies for arterial corridors
Speeding	Positive effects	Further research into crash sequence, enforcement strategies.
Drunk driving	Negative effects	Enforcement strategies for arterial corridors
Aggressive driving	Greater negative effects than drunk driving	Enforcement strategies for arterial corridors, including time of day
Type of vehicle	Van, LTV's and pickups with positive effects compared to passenger cars	Road design characteristics to improve passenger car visibility of signals and lane assignments. Future research of driver age and gender relationships with type of vehicle used.

Similarities regarding injury severity may be more closely related to crash mechanism and land use than to the location of the crash for arterial corridors. This does not necessarily contradicts the differences between road locations, but points to some additional correlations between closely spaced access points, major intersections, traffic congestion and driveway interactions by land use. The intersection crash classification used in the crash database may also have an effect on these systematic results, as suggested by Wang et al. (2008). This investigation has provided a glimpse of the systematic trends in regards to injury severity on arterial corridors. The performance of the joint models pointed to some contributing factors that only showed effects on intersections or segment models, but also significance in the overall model. These factors included *adt* per lane, access class, lane width, lighting density and skid resistance number. Their safety effects and broader implications are in Table 6-2, page 275.

Table 6-2: Safety Effects and Implications of the Most Important Roadway-related Contributing Factors

Contributing Factor	Safety Effects	Implications
Speed limit	Positive effects of arterials with lower speed limits (40-45 mph)	Possible drawbacks of higher speed arterials. Future research on relationships between speed limits 40-45, driveway density and land use.
***adt* per lane**	Positive effects of increased average traffic density, not significant at signalized int	Traffic management strategies should take into account effects on operating speeds to avoid negative safety effects.
Access Class	Positive effects of classes 5-7 (vs. 2-4) in signal, overall, urban area models	Benefits of access management strategies in urban areas (preliminary result). Future research need to stratify class 5 by speed limit.
Land Use	Positive effects of urban areas	Rural area road design characteristics future research. Driveway research limited to urban areas.
Lane width	Positive effect of wide (>12 ft) lanes in urban, narrow (<11 ft) in rural areas. Entity models showed positive effects at intersections.	Design of multilane arterial lane widths varying according to land use. Bicycle lane separation may have negative effects on motor vehicle users. Positive effects at intersections may point to limited application to intersection influence area. Further research is needed.
Sidewalk width	Moderate to major positive effects on intersections and rural area model	Further proof of safety benefits of pedestrian facilities to drivers. Rural area finding is preliminary, further research is needed.
Non-high mast Lighting density	Partial non-high mast density had positive effects in segments, rural areas and rear-end crashes.	Points to the possible design treatments in rural areas that may affect the land use effect discussed earlier. Crash type models point to systematic countermeasure.
High mast Lighting density	Negative effects of total or partial high mast density	Investigation into high-mast lighting locations (i.e. interchanges) to recommend systematic treatments and/or design changes
Type Friction Course	Positive effects of newer friction courses in most models and urban area	Points to possible systematic benefits of resurfacing in arterial corridors. Old friction courses become ineffective. Urban area tends to support larger traffic and pavement wear.
Skid Resistance	Negative effects when comparing current cut-off value of 35. Positive urban area effect.	Preliminary analysis indicated a positive effect of skid numbers >44 in urban areas, but negative effects in rural areas. Earlier studies point to driver behavior offset (Elvik and Greibe, 2005)
Median	Not significant	Median size or type not as important in design as spacing, further research is required
Number of approaches	Not significant	Intersection size rather than the number of approaches seems to have safety effects

With the information provided by these models (refer to Table 6-2, page 275), there is enough evidence to focus additional investigation and evaluation of certain locations, such as those with high-mast lighting. On the other hand, some of the contributing factors point to systemic effects that can be used in improving road design given results from additional research. In addition, some present strategies, such as the bicycle lane separation may have a negative effect in urban areas, were the results from this investigation suggest that wider lanes are desirable. Due to the macro approach of this research, the exact nature of some road related-effects is not completely described in the models. However, the potential of the joint analysis for future research dealing with arterial corridors was demonstrated. Also, there were indications that this analysis can be improved if it is stratified by land use. There are some caveats to this approach that can be better handled if a comprehensive analysis with models by crash type and road entity are compared.

While this investigation attempted a large scale approach to the analysis of arterial corridors, some of the lessons learned here could be applied to the analysis of individual arterial corridors. The application of systematic countermeasures based on large scale could be useful when analyzing individual corridors. This strategy has been applied in jurisdictions with safe corridor programs, such as Virginia. Their approach includes a regional analysis determined by crash trends, socioeconomic and geographical characteristics. Future large scale analysis with a representative sample of arterial corridors can provide results that confirm or extend the ones obtained in this investigation.

6.7 Recommendations for Further Research

The results of this investigation are useful, but limited to the injury severity analysis for one person involved in each crash. An extension of this research using statistical methods that account for the correlation between involvements in the same crash should be a research priority. These models will take full advantage of the wide data format that could not be utilized in the final analysis. The use of the wide format was nonetheless very useful and is recommended in future research work that includes severity analysis. In addition, separate analysis of single and multiple vehicle crash involvements with their different crash mechanisms and their relationships with road characteristics is recommended.

Additional research into the relationships of crashes and road features in urban and rural areas is needed. Additional comparison between models by land use, road entity and crash types would confirm whether the land use approach is best for all injury severity analyses of high-speed multilane arterials. Crashes in rural and urban areas are affected by some common road characteristics that follows design standards or guidelines, for example geometric (lane width) and pavement (friction course). The internal correlations between these characteristics and the driver injury severity outcomes are of interest. Also, to review and to possibly use additional data from the most recent years (2005-2007) to evaluate how some road characteristics, such as friction courses, have changed over time in sufficient quantities for performing additional systematic crash frequency and severity analysis; this may allow investigating the safety effectiveness of newer friction course mixes not widely available in 2002-2004 and comparing these results to earlier analysis with older crash data.

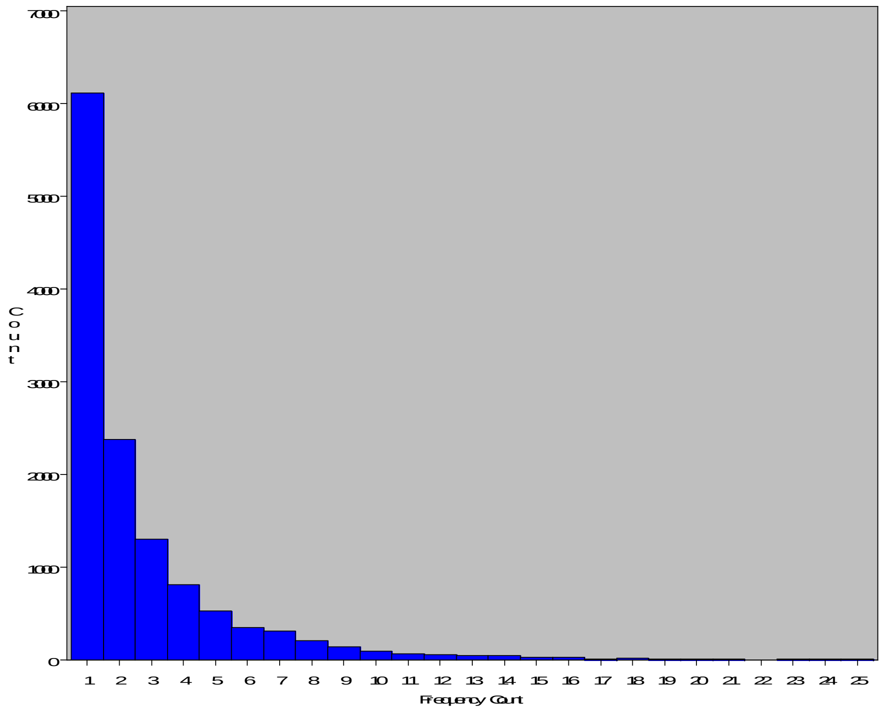
Another area of interest in injury severity analysis is the relationship between the ejection event and seat belt use. Even tough this event is technically a post-crash event; its inclusion in the model might have measured some unobserved factors. In addition, recent studies have pursued the use of statistical method that account for the over reporting of seat belt usage in minor crashes.

Preliminary analysis suggested that there are no significant differences between the severe crashes in state vs. non-state roads for the high-speed multilane arterials. Although any arterial corridor, even at the local level will likely comply with the minimum standards and guidelines, research needs to address other possibilities. Differences between state and non-state roads could include access management strategies, lack of uniform guidelines within a county (by cities) or region (by counties), land uses, law enforcement, and travel choices. There are efforts underway to include more local (county and city) road and crash data in an integrated crash database. This will allow further research including non-state roads in the future.

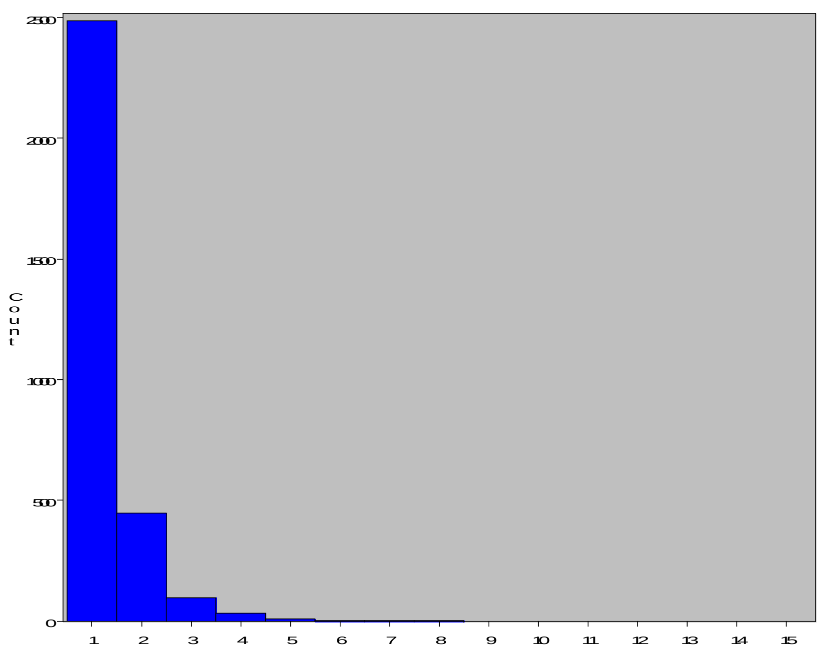
APPENDIX A: INTERSECTIONS WITH TOTAL AND SEVERE CRASH COUNTS FOR THE YEAR 2004

Distribution of Total and Severe Crash Counts (All Intersections in Multilane Arterials)

Distribution of Total Crashes At (or Near) All Intersections in Multilane Arterials

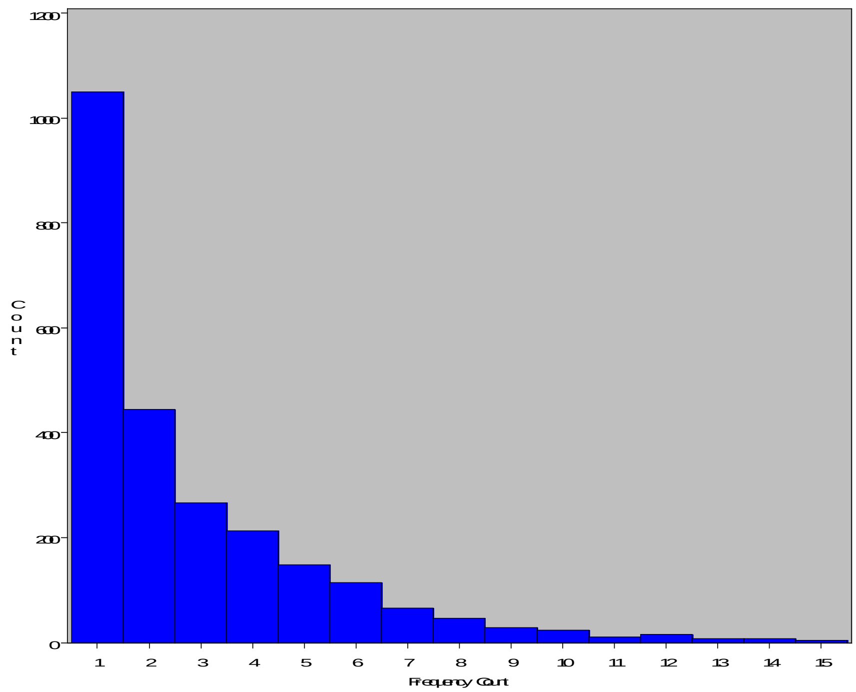


Severe Crashes At (or Near) All Intersections in Multilane Arterials

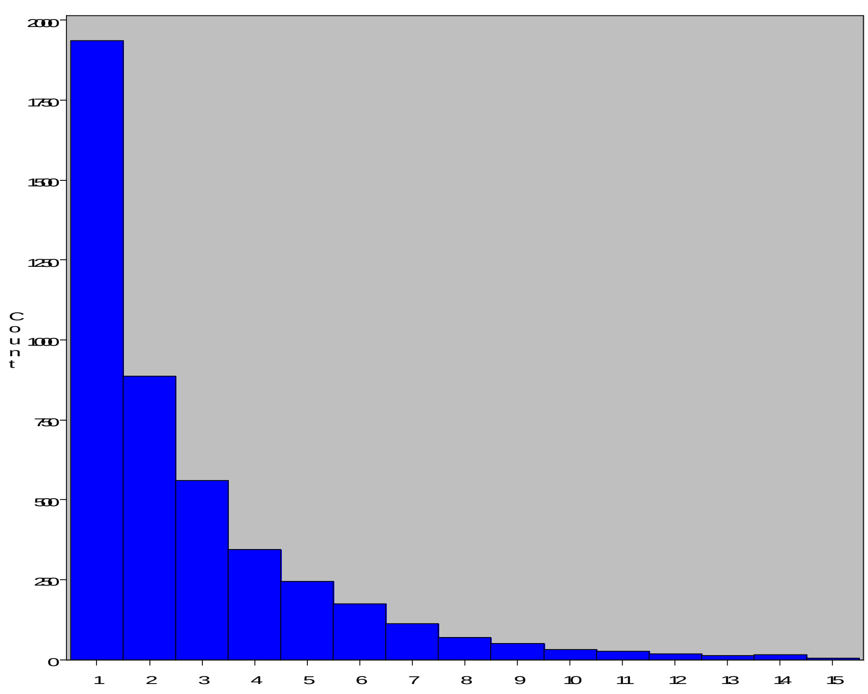


Distribution of Total Crash Counts (Signalized Intersections in Rural and Urban Multilane Arterials)

Total Crashes At (or Near) Signalized Rural Intersections in Multilane Arterials

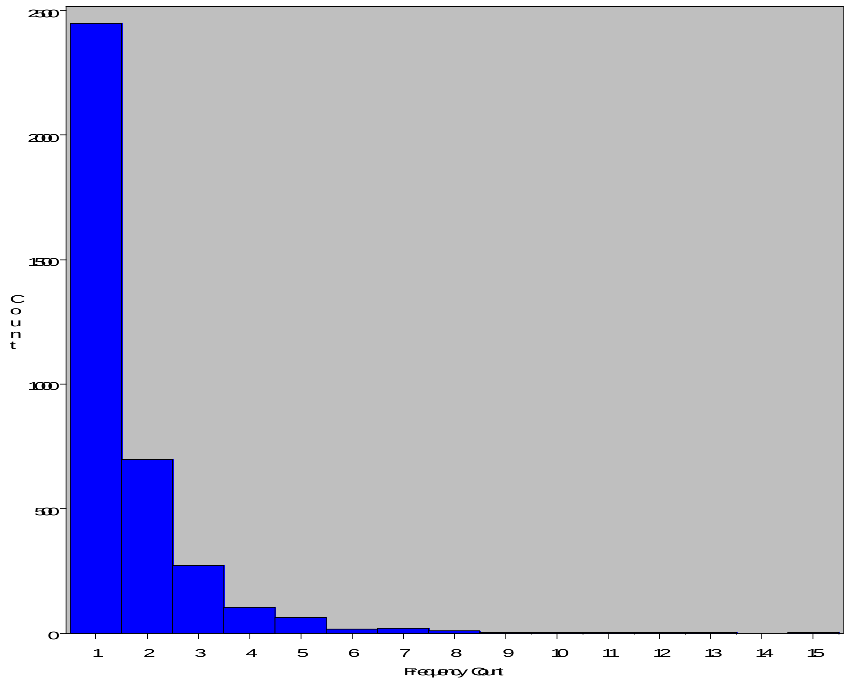


Total Crashes At (or Near) Signalized Urban Intersections in Multilane Arterials

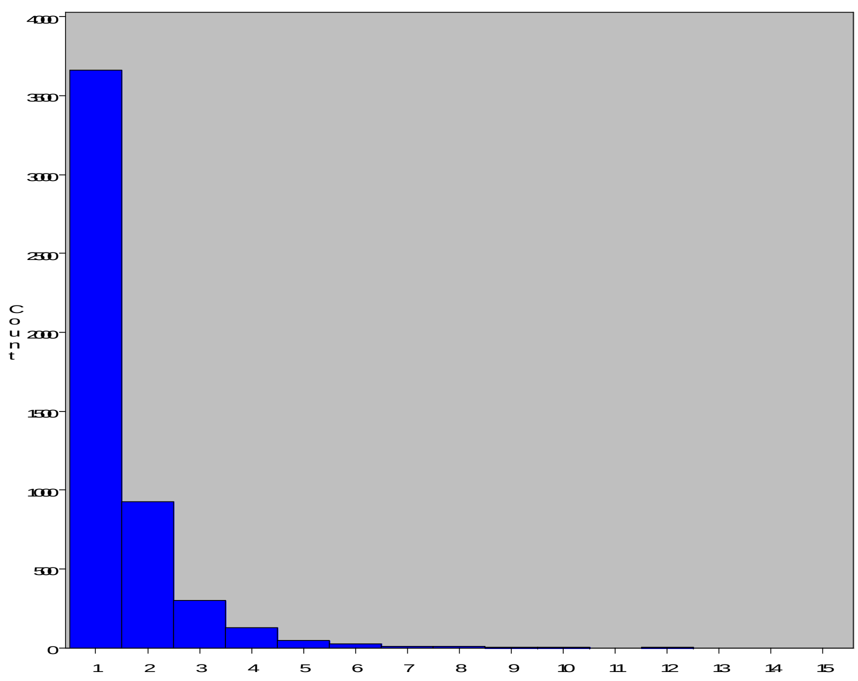


Distribution of Total Crash Counts (Unsignalized Intersections in Rural and Urban Multilane Arterials)

Total Crashes At (or Near) Unsignalized Rural Intersections in Multilane Arterials

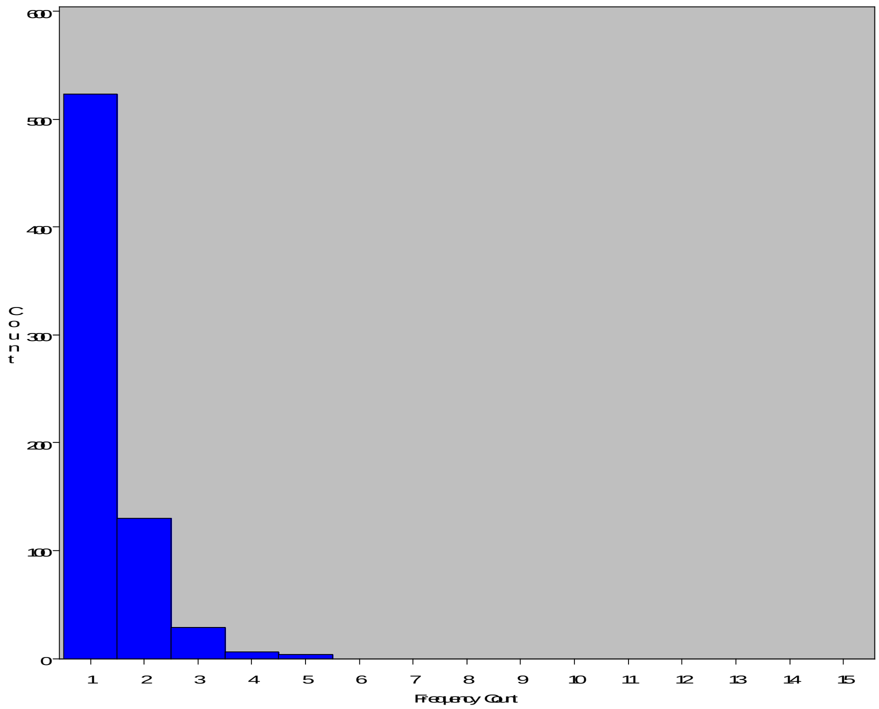


Total Crashes At (or Near) Unsignalized Urban Intersections in Multilane Arterials

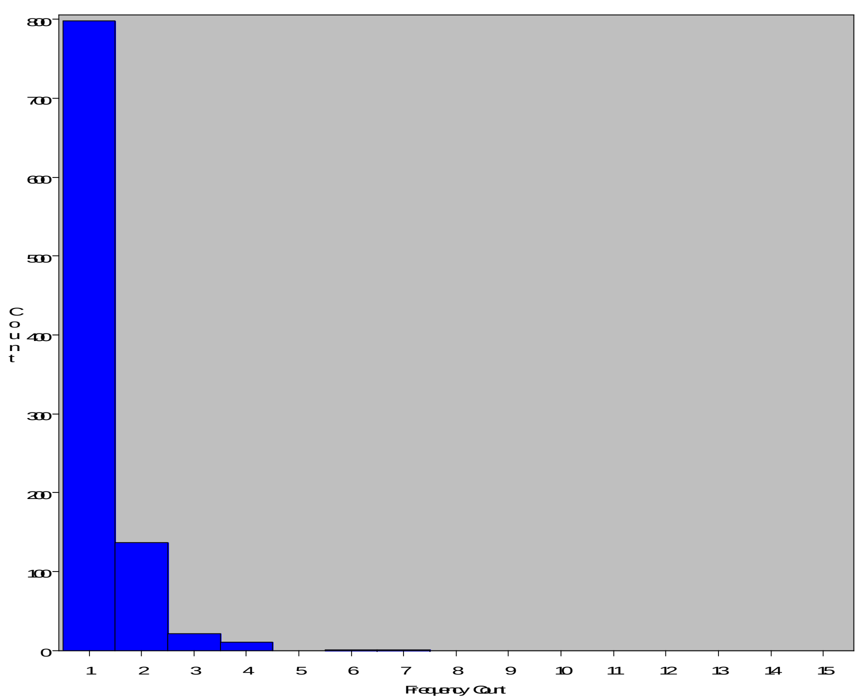


Distribution of Severe Crash Counts (Signalized Intersections in Rural and Urban Multilane Arterials)

Severe Crashes At (or Near) Signalized Rural Intersections in Multilane Arterials

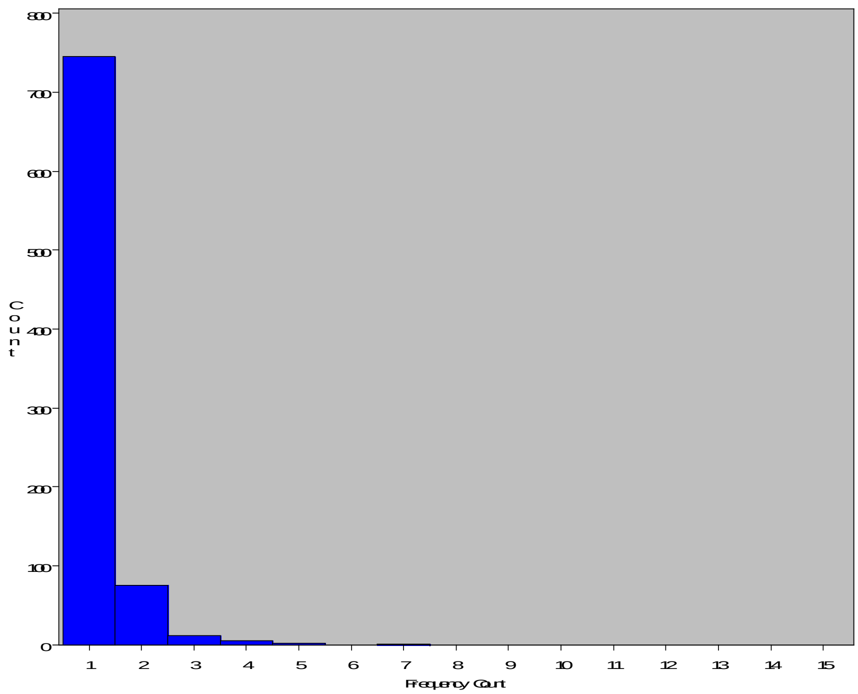


Severe Crashes At (or Near) Signalized Urban Intersections in Multilane Arterials

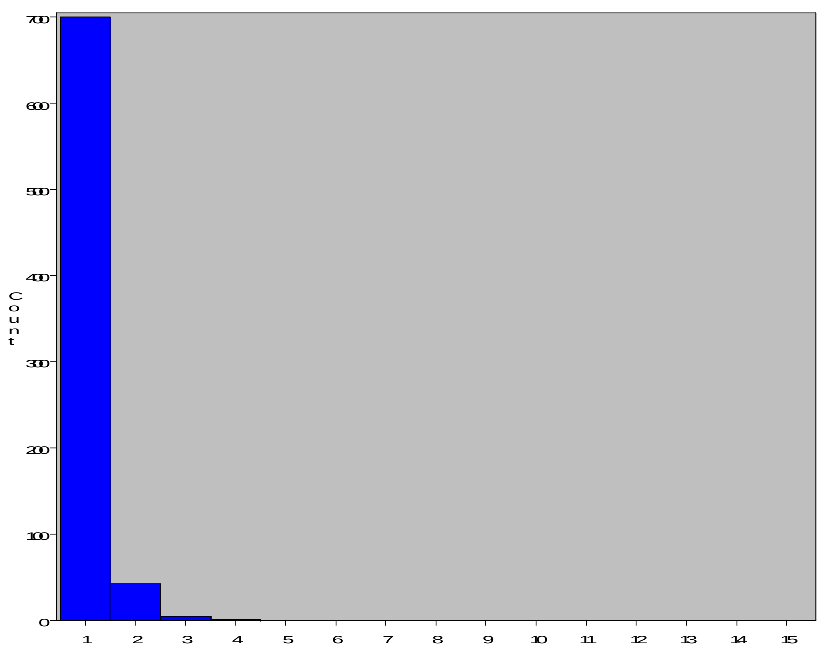


Distribution of Severe Crash Counts (Unsignalized Intersections in Rural and Urban Multilane Arterials)

Severe Crashes At (or Near) Unsignalized Rural Intersections in Multilane Arterials



Severe Crashes At (or Near) Unsignalized Urban Intersections in Multilane Arterials



APPENDIX B: CATEGORICAL DATA'S FINAL ANALYSIS

Categorical Data Analysis (All High-speed Multilane Involvements Records; N=215,898)

		Driver 1			Driver 2		
Variable	DF PCHI	p-value	CONTGY	CRAM V	p-value	CONTGY	CRAM V
Year	2	0.0647	0.0054	0.0054	0.0171	0.0065	0.0065
Driver_Ageg_Group1	4	<0.0001	0.0411	0.0412	<0.0001	0.0174	0.0174
nSex1	2	<0.0001	0.0232	0.0232	<0.0001	0.0185	0.0185
nFirst_Safety_Equipment1	7	<0.0001	0.2035	0.2079	<0.0001	0.1639	0.1661
Speeding1	2	<0.0001	0.0959	0.0963	<0.0001	0.0674	0.0676
nFirst_Contributing_Cause1	25	<0.0001	0.1143	0.1150	<0.0001	0.0578	0.0579
nVehicle_Fault_Code1	1	<0.0001	0.0739	0.0741	<0.0001	0.0198	0.0198
Red_light_running1	1	<0.0001	0.0215	0.0215	<0.0001	0.0098	0.0098
nResidence_Code1	5	<0.0001	0.0294	0.0294	<0.0001	0.0176	0.0176
nPhysical_Defects1	7	<0.0001	0.0560	0.0561	<0.0001	0.0167	0.0167
nEjected1	3	<0.0001	0.2287	0.2349	<0.0001	0.1831	0.1862
nRecommend_Re_Exam1	2	<0.0001	0.0228	0.0228	0.5028	0.0027	0.0027
nRace1	4	<0.0001	0.0295	0.0295	0.0006	0.0101	0.0101
nFirst_Harmful_Event1	39	<0.0001	0.1719	0.1745	<0.0001	0.1034	0.1039
nOn_Off_Roadway	1	<0.0001	0.0796	0.0798	0.0004	0.0081	-0.0081
nPoint_of_Impact1	21	<0.0001	0.1490	0.1507	<0.0001	0.0997	0.1002
nVehicle_Movement1	13	<0.0001	0.0687	0.0689	<0.0001	0.0621	0.0622
nType_of_Vehicle1	15	<0.0001	0.1811	0.1842	<0.0001	0.1794	0.1824
nVehicle_Use1	16	<0.0001	0.0407	0.0408	<0.0001	0.0331	0.0331
CRASH_LANE	7	<0.0001	0.0650	0.0652	<0.0001	0.0373	0.0373
nRural_Urban	1	<0.0001	0.0442	-0.0442	<0.0001	0.0285	-0.0285
nLocation_Type	2	<0.0001	0.0747	0.0749	<0.0001	0.0341	0.0341
nVehicle_Special _Functions1	6	<0.0001	0.0193	0.0193	<0.0001	0.0145	0.0145
nFirst_Vehicle_Defect1	9	<0.0001	0.0378	0.0378	<0.0001	0.0196	0.0196
nCrash_Fault_Code	1	<0.0001	0.0819	0.0821	<0.0001	0.0466	-0.0467
nTotal_Number_of_Drivers	11	<0.0001	0.1212	0.1221	<0.0001	0.0261	0.0261
nWork_Area1	3	0.0143	0.0075	0.0075	0.0011	0.0084	0.0084
nLocation_on_Roadway1	5	<0.0001	0.0991	0.0996	<0.0001	0.0130	0.0130
nAlcohol_Drug_Use1	6	<0.0001	0.1034	0.1039	<0.0001	0.0574	0.0575
nDivided_Undivided _Highway	2	<0.0001	0.0207	0.0207	<0.0001	0.0137	0.0137
nTrafficway_Character	3	<0.0001	0.0608	0.0609	<0.0001	0.0191	0.0191
nType_of_Shoulder	2	<0.0001	0.0295	0.0295	<0.0001	0.0160	0.0160
nRoad_Surface_Type	5	<0.0001	0.0162	0.0162	<0.0001	0.0157	0.0157
nNumber_of_Lanes	28	<0.0001	0.0432	0.0433	<0.0001	0.0239	0.0239
nCRRATECD	32	<0.0001	0.0997	0.1002	<0.0001	0.0521	0.0521
Median_type	2	<0.0001	0.0201	0.0201	0.0022	0.0095	0.0095
nFirst_Traffic_Control	12	<0.0001	0.0513	0.0514	<0.0001	0.0468	0.0468
nSite_Location	3	<0.0001	0.0517	0.0518	<0.0001	0.0601	0.0602
nFirst_Road_Condition	9	0.6691	0.0060	0.0060	0.5197	0.0065	0.0065
nTYPEPARK	7	<0.0001	0.0512	0.0512	<0.0001	0.0328	0.0328

		Driver 1			Driver 2		
Variable	DF PCHI	p-value	CONTGY	CRAM V	p-value	CONTGY	CRAM V
nNUM_LEGS	3	<0.0001	0.0214	0.0214	0.0006	0.0145	0.0145
nRDACCESS	1	0.7822	0.0008	0.0008	0.0379	0.0056	0.0056
TIME_GROUP	3	<0.0001	0.0430	0.0430	<0.0001	0.0157	0.0158
nLighting_Condition	5	<0.0001	0.0595	0.0596	<0.0001	0.0267	0.0267
nWeather	5	<0.0001	0.0209	0.0209	<0.0001	0.0178	0.0178
nDay_of_Week	6	<0.0001	0.0134	0.0134	0.9511	0.0029	0.0029
nRoad_Surface_Condition	4	<0.0001	0.0213	0.0213	<0.0001	0.0180	0.0180
nFirst_Vision_Obstructed	10	0.6020	0.0066	0.0066	<0.0001	0.0189	0.0189
month	11	0.0085	0.0116	0.0116	0.3190	0.0081	0.0081

NOTES:

1) DF_PCHI= Degrees of freedom of independence Pearson Chi test, CONTGY= Contingency coefficient, CRAM V= Cramer's V coefficient

2) Horizontal lines separate the three main variable groups in the following order: driver-, crash-vehicle-, road- and environment-related.

3) Variables in shadow are statistically independent from the severe driver injury (response) for at least one of the drivers in the analysis.

Categorical Data Analysis (High-speed Multilane Involvements with Complete Records; N=156,688)

Variable	DF PCHI	Driver 1 p-value	Driver 1 CONTGY	Driver 1 CRAM V	Driver 2 p-value	Driver 2 CONTGY	Driver 2 CRAM V
Year	2	0.0589	0.0065	0.0065	0.0055	0.0086	0.0086
Driver_Ageg_Group1	4	<0.0001	0.0419	0.0420	<0.0001	0.0168	0.0168
nSex1	2	<0.0001	0.0229	0.0229	<0.0001	0.0146	0.0146
nFirst_Safety_Equipment1	7	<0.0001	0.2066	0.2111	<0.0001	0.1674	0.1698
Speeding1	2	<0.0001	0.0963	0.0967	<0.0001	0.0664	0.0666
nFirst_Contributing_Cause1	25	<0.0001	0.1129	0.1136	<0.0001	0.0543	0.0544
nVehicle_Fault_Code1	1	<0.0001	0.0789	0.0792	<0.0001	0.0168	0.0168
Red_light_running1	1	<0.0001	0.0195	0.0195	0.0001	0.0102	0.0102
nResidence_Code1	5	<0.0001	0.0290	0.0290	0.0002	0.0132	0.0132
nPhysical_Defects1	7	<0.0001	0.0559	0.0560	<0.0001	0.0157	0.0157
nEjected1	3	<0.0001	0.2355	0.2424	<0.0001	0.1891	0.1926
nRecommend_Re_Exam1	2	<0.0001	0.0247	0.0247	0.4068	0.0036	0.0036
nRace1	4	<0.0001	0.0312	0.0312	0.0014	0.0112	0.0112
nFirst_Harmful_Event1	39	<0.0001	0.1758	0.1786	<0.0001	0.1096	0.1103
nOn_Off_Roadway	1	<0.0001	0.0768	0.0770	0.0472	0.0053	-0.0053
nPoint_of_Impact1	21	<0.0001	0.1558	0.1577	<0.0001	0.1021	0.1027
nVehicle_Movement1	13	<0.0001	0.0674	0.0676	<0.0001	0.0642	0.0643
nType_of_Vehicle1	15	<0.0001	0.1827	0.1858	<0.0001	0.1848	0.1881
nVehicle_Use1	16	<0.0001	0.0418	0.0419	<0.0001	0.0333	0.0333
CRASH_LANE	7	<0.0001	0.0988	0.0993	<0.0001	0.0369	0.0369
nRural_Urban	1	<0.0001	0.0587	-0.0588	<0.0001	0.0389	-0.0389
nLocation_Type	2	<0.0001	0.0860	0.0863	<0.0001	0.0397	0.0397
nVehicle_Special_Functions1	6	<0.0001	0.0203	0.0203	<0.0001	0.0157	0.0157
nFirst_Vehicle_Defect1	9	<0.0001	0.0419	0.0419	<0.0001	0.0187	0.0187
nCrash_Fault_Code	1	<0.0001	0.0845	0.0849	<0.0001	0.0471	-0.0471
nTotal_Number_of_Drivers	11	<0.0001	0.1200	0.1209	<0.0001	0.0283	0.0283
nWork_Area1	3	0.0893	0.0069	0.0069	0.0211	0.0074	0.0074
nLocation_on_Roadway1	5	<0.0001	0.0981	0.0986	<0.0001	0.0140	0.0140
nAlcohol_Drug_Use1	6	<0.0001	0.1052	0.1058	<0.0001	0.0622	0.0623
nDivided_Undivided _Highway	2	<0.0001	0.0191	0.0191	<0.0001	0.0124	0.0124
nTrafficway_Character	3	<0.0001	0.0579	0.0580	<0.0001	0.0203	0.0203
nType_of_Shoulder	2	<0.0001	0.0397	0.0397	<0.0001	0.0189	0.0189
nRoad_Surface_Type	5	<0.0001	0.0159	0.0159	0.0004	0.0128	0.0128
nNumber_of_Lanes	28	<0.0001	0.0470	0.0471	0.0003	0.0199	0.0199
nCRRATECD	32	<0.0001	0.0997	0.1002	<0.0001	0.0521	0.0521
Median_type	2	<0.0001	0.0201	0.0201	0.0022	0.0095	0.0095
nFirst_Traffic_Control	12	<0.0001	0.0554	0.0554	<0.0001	0.0434	0.0435
nSite_Location	3	<0.0001	0.0514	0.0515	<0.0001	0.0585	0.0586
nFirst_Road_Condition	9	0.1266	0.0101	0.0101	0.6185	0.0072	0.0072
nTYPEPARK	7	<0.0001	0.0512	0.0512	<0.0001	0.0328	0.0328
nNUM_LEGS	3	<0.0001	0.0214	0.0214	0.0006	0.0145	0.0145

Variable	DF PCHI	Driver 1			Driver 2		
		p-value	CONTGY	CRAM V	p-value	CONTGY	CRAM V
nRDACCESS	1	0.7822	0.0008	0.0008	0.0379	0.0056	0.0056
TIME_GROUP	3	<0.0001	0.0429	0.0429	<0.0001	0.0163	0.0163
nLighting_Condition	5	<0.0001	0.0703	0.0704	<0.0001	0.0339	0.0339
nWeather	5	<0.0001	0.0220	0.0220	<0.0001	0.0151	0.0151
nDay_of_Week	6	<0.0001	0.0157	0.0157	0.9382	0.0036	0.0036
nRoad_Surface_Condition	4	<0.0001	0.0212	0.0212	<0.0001	0.0141	0.0141
nFirst_Vision_Obstructed	10	0.4711	0.0084	0.0084	<0.0001	0.0174	0.0174
month	11	0.0143	0.0132	0.0132	0.1237	0.0108	0.0108

NOTES:

1) DF_PCHI= Degrees of freedom of independence Pearson Chi test, CONTGY= Contingency coefficient, CRAM V= Cramer's V coefficient

2) Horizontal lines separate the three main variable groups in the following order: driver, crash-vehicle, road and environment-related.

3) Variables in shadow are statistically independent from the severe driver injury (response) for at least one of the drivers in the analysis.

Test of Independence between Driver Section Number and the Variables Listed (Sample n=118,790; Complete N= 394,394)

Variable	p-value using random sample	p-value using complete sample
Severe_driver_x	0.000298199	0.0002982
Year	0.000298199	1
Driver_Ageg_Group_x	0.177952724	<0.0001
Gender_x	<0.0001	<0.0001
Safety_Equipment_x	<0.0001	<0.0001
Speeding_x	<0.0001	<0.0001
Contributing_Cause_x	<0.0001	<0.0001
At_Fault_driver_x	<0.0001	<0.0001
Red_light_running_x	<0.0001	<0.0001
Residence_Code_x	<0.0001	0.00613153
Physical_Defects_x	0.325869667	<0.0001
Ejected_x	<0.0001	<0.0001
nRecommend_Re_Exam_x	<0.0001	<0.0001
nRace_x	<0.0001	<0.0001
Harmful_Event_Group_x	<0.0001	<0.0001
Off_Roadway	<0.0001	1
Point_Impact_x	<0.0001	<0.0001
Vehicle_Maneuver_x	<0.0001	<0.0001
Type_of_Vehicle_x	<0.0001	<0.0001
Private_vehicle_use_x	<0.0001	<0.0001
CRASH_LANE5	0.003356631	1
nRural_Urban	<0.0001	1
Location_Type	0.751214189	1
nVehicle_Special_Functions_x	0.00312239	<0.0001
nFirst_Vehicle_Defect_x	<0.0001	<0.0001
nCrash_Fault_Code	<0.0001	1
nTotal_Number_of_Drivers	<0.0001	1
nWork_Area_x	0.021756961	0.49171805
nAlcohol_Drug_Use_x	0.397420419	<0.0001
Undivided_Highway	<0.0001	1
Roadway_Curve	0.082809347	1
nType_of_Shoulder	0.000492899	1
Concrete_Surface	<0.0001	1
Number_of_Lanes	0.625110641	1
nCRRATECD	0.045392449	1
Median_type	0.000126276	1
Traffic_Control	0.171939914	1
Intersection	<0.0001	1
nFirst_Road_Condition	<0.0001	1
Speed_limit_x	0.602929272	<0.0001

Variable	p-value using random sample	p-value using complete sample
nTYPEPARK	<0.0001	1
nNUM_LEGS	0.22404173	1
nRDACCESS	0.105439642	1
TIME_GROUP2	<0.0001	1
Lighting_Condition	<0.0001	1
Weather	0.988194757	1
Day_of_Week	0.259111505	1
nRoad_Surface_Condition	0.192308862	1
nFirst_Vision_Obstructed	<0.0001	1
RAIN_SEASON	0.320956912	1

APPENDIX C: LIST OF VARIABLES CONSIDERED IN FINAL ANALYSIS

Variables Considered in Final Analysis
(Variables in **bold** were added from the RCI data merge)

Type	Variable	Value	Description
DRIVER-RELATED VARIABLES	Driver_Ageg_Group_x	1	25-64 years
		2	15-19 years
		3	20-24 years
		4	65-79 years
		5	80-98 years
	Gender_x	0	Male
		1	Female
	Safety_Equipment_x	0	None or other
		1	Seat belt / child restraint
	Speeding_x	0	Not Speeding
		1	Speeding
		2	Unknown
	Contributing_Cause_x	1	No improper driver action
		2	Aggressive Driving
		3	Alcohol / Drugs
		4	Other
	At_Fault_driver_x	0	Not cited
		1	Cited
	Red_light_running_x	0	No
		1	Yes
	Residence_Code_x	0	Not Florida
		1	Florida Resident
	Physical_Defects_x	0	No
		1	Yes
	Ejected_x	0	No
		1	Total or partial
CRASH	Harmful_Event_Group_x	1	Rear-End
		2	Head-On
		3	Angle
		4	Left Turn
		5	Sideswipe
		6	Fixed Object
		7	Other
	Off_Roadway	0	No
		1	Yes
	Point_Impact_x	0	Not Driver's side
		1	Driver's side

** The first level shown is the base value for the coefficients

Variables Considered in Final Analysis (Continued)
(Variables in **bold** were added from the RCI data merge)

Type	Variable	Value	Description
VEH-COLLISION VARIABLES	Vehicle_Maneuver_x	1	Straight Ahead
		2	Slowing / Stopping
		3	Left Turn
		4	Other
	Type_of_Vehicle_x	1	Automobile
		2	Van, Light Truck, Pick up
		3	Trucks and buses
		4	Bicycle and motorcycle
		5	Other
	Private_vehicle_use_x	0	No
		1	Yes
	nRural_Urban	0	Rural
		1	Urban
	Location_Type	1	Residential
		2	Business
		3	Open Country
	Multivehicle	0	No
		1	Yes
	Vehicle_Defect	0	No
		1	Yes
	nWork_Area_x	1	None
		2	Nearby
		3	Entered
ROAD-RELATED	Undivided_Highway	0	No
		1	Yes
	Roadway_Curve	0	No
		1	Yes
	Concrete_Surface	0	No
		1	Yes
	Number_of_Lanes	4	4 lane major road
		5	5 lane major road
		6	6 lane major road
		7	7 lane major road
		8	8 lane major road
	Median_type	1	Raised
		2	Paved

** The first level shown is the base value for the coefficients

Variables Considered in Final Analysis

(Variables in **bold** were added from the RCI data merge) Continued

Type	Variable	Value	Description
ROADWAY-RELATED VARIABLES	Traffic_Control	1 2 3	Other control or none Traffic signal or yield Stop sign or flashing lights
	Intersection	0 1	No Yes
	Speed_limit_x	0 1	Less than 40 mph, more than 45 mph 40-45 mph
	nType_of_Shoulder	1 2 3	Paved Unpaved Curb
	Road_Condition	1 2 3	Dry Wet Other
	ADT_PER_LANE	continuous	Section *aadt* / Number of lanes
	Median_size	1 2 3 4	Median width ≥ 40 ft Median width < 15.5 ft 15.5 ft ≤ Median width < 19.5 ft 19.5 ft ≤ Median width < 40 ft
	nAVGTFACT	continuous	Percentage of *aadt* that consists of trucks
	Skid_Resistance	0 1	1 ≤ Friction Number < 35 Friction Number ≥ 35
	Lane_width	1 2 3 4	11 ft ≤ Lane width ≤ 12 ft Lane width < 10 ft 10 ft ≤ Lane width < 11 ft Lane width > 12 ft
	Shoulder_width (outside shoulder)	1 2 3 4	Shoulder width < 6 ft 6 ft ≤ Shoulder width < 8 ft 8 ft ≤ Shoulder width < 10 ft Shoulder width ≥ 10 ft
	LIGHTCDE	N P Y	One or none non-high mast lighting exists Partial non-high mast lighting exists (Rates of 4-24 lights per mile) Full non-high mast lighting exists (Rates of 25 lights per mile or more)

** The first level shown is the base value for the coefficients

Variables Considered in Final Analysis (Continued)

(Variables in **bold** were added from the RCI data merge)

Type	Variable	Value	Description
ROADWAY-RELATED VARIABLES	**LIGHTING**	N	One or none high mast lighting exists
		P	Partial high mast lighting exists (Rates of 4-9 lights per mile)
		Y	Full high mast lighting exists (Rates of 10 lights per mile or more)
	Access_class	2	Access Management Class 2,3,4
		5	Access Management Class 5
		6	Access Management Class 6
		7	Access Management Class 7
		9	Non applicable
	AUX_Lane_Type	1	No auxiliary lanes (base value)
		9	Left Turn lane (others)
		9	Right Turn lane (others)
		9	Bus lane (others)
		5	Merging (inside)
		6	Merging (outside)
		7	Parking lanes
	AUX_Lane_Num	1	No auxiliary lanes
		2	One auxiliary lanes
		3	Two auxiliary lanes
		4	Three auxiliary lanes
		5	More than three auxiliary lanes
	Type_Friction_Course	0	Friction course type 2 (FC-2)
		1	Friction course type 1 (FC-1)
		4	Friction course type 4 (FC-4)
		5	Friction course type 5 (FC-5)
		9	Not applicable or other (FC-3, 6)
	Pavement_condition	1	Poor or very poor
		2	Fair
		3	Good
		4	Very good
	Sidewalk_width_group	1	Sidewalk width < 4 ft
		2	4 ft ≤ Sidewalk width < 6 ft
		3	Sidewalk width ≥ 6 ft

** The first level shown is the base value for the coefficients

Variables Considered in Final Analysis (Continued)
(Variables in **bold** were added from the RCI data merge)

Type	Variable	Value	Description
ROAD	**Urban_size**	1	Rural
		2	Small Urban or urbanized
		3	Large Urbanized or Metropolitan
ENVIRONMENTAL	TIME_GROUP2	0	Time of crash: 6AM- 6PM
		1	Time of crash: 6PM- 6AM
	Lighting_Condition	1	Daylight / Dusk / Dawn
		2	Dark with street lighting
		3	Dark without street lighting
	Road_Surface_Condition	1	Dry
		2	Wet
		3	Slippery or Icy
		4	Other
	Vision_Obstructed	1	Vision Not Obscured
		2	Inclement Weather, Fog, Smoke, Glare
		3	Parked/Stopped Vehicle
		4	Other

** The first level shown is the base value for the coefficients

APPENDIX D: INJURY SEVERITY MODELS' FINAL ANALYSIS – MODELS BY ROAD ENTITY

Injury Severity Regression Model
(for a Sample of Driver Involvements of All Crashes)

Parameter		DF	Estimate	Standard Error	Wald chi-square	Pr > chi-square
Intercept		1	-0.8615	0.1441	35.7257	<.0001
Driver_Ageg_Group_x	5	1	0.4829	0.077	39.3597	<.0001
Driver_Ageg_Group_x	4	1	0.3523	0.048	53.8723	<.0001
Driver_Ageg_Group_x	3	1	-0.2494	0.0424	34.692	<.0001
Driver_Ageg_Group_x	2	1	-0.2648	0.0471	31.6077	<.0001
Ejected_x		1	1.4517	0.0693	438.965	<.0001
Speeding_x	2	1	-0.1468	0.1202	1.4914	0.222
Speeding_x	1	1	-0.894	0.0918	94.7695	<.0001
Gender_x		1	0.1963	0.0631	9.6756	0.0019
Safety_Equipment_x	1	1	-1.1936	0.0465	660.1324	<.0001
Gender_x*Safety_Equi	1	1	0.2195	0.0705	9.6942	0.0018
At_Fault_driver_x		1	-0.6207	0.0387	257.5631	<.0001
Residence_Code_x		1	0.1611	0.0675	5.6861	0.0171
Physical_Defects_x		1	0.4286	0.0743	33.2722	<.0001
Harmful_Event_Group_	7	1	0.0928	0.0438	4.4981	0.0339
Harmful_Event_Group_	6	1	0.5934	0.07	71.8379	<.0001
Harmful_Event_Group_	5	1	-0.2495	0.0898	7.7106	0.0055
Harmful_Event_Group_	4	1	0.8073	0.0597	182.945	<.0001
Harmful_Event_Group_	3	1	0.5789	0.0469	152.2623	<.0001
Harmful_Event_Group_	2	1	1.056	0.084	157.8853	<.0001
Contributing_Cause_x	4	1	0.4732	0.04	140.2797	<.0001
Contributing_Cause_x	3	1	0.5583	0.1108	25.3682	<.0001
Contributing_Cause_x	2	1	0.4655	0.0496	88.0634	<.0001
Type_of_Vehicle_x	5	1	-0.2796	0.1371	4.1565	0.0415
Type_of_Vehicle_x	4	1	0.0491	0.0784	0.3915	0.5315
Type_of_Vehicle_x	3	1	-1.0309	0.113	83.1693	<.0001
Type_of_Vehicle_x	2	1	-0.1983	0.035	32.1383	<.0001
point_impact_x		1	0.0874	0.0725	1.4552	0.2277
point_imp*Speeding_x	2	1	0.2149	0.1349	2.5378	0.1111
point_imp*Speeding_x	1	1	0.3451	0.0889	15.0605	0.0001
Off_Roadway		1	-0.4899	0.0798	37.6589	<.0001
Off_Roadw*Speeding_x	2	1	-0.2443	0.1266	3.7237	0.0536
Off_Roadw*Speeding_x	1	1	0.2539	0.0972	6.8214	0.009
Off_Roadw*Multivehic	1	1	0.7095	0.1195	35.2598	<.0001

Parameter		DF	Estimate	Standard Error	Wald chi-square	Pr > chi-square
nWork_Area_x	3	1	-0.1907	0.1037	3.3812	0.0659
nWork_Area_x	2	1	-0.2875	0.0887	10.5022	0.0012
Multivehicle	1	1	-0.7576	0.1088	48.4819	<.0001
Intersect*Multivehic	1	1	0.3894	0.0984	15.6573	<.0001
Speed_limit_x	1	1	-0.3917	0.0354	122.3589	<.0001
ADT_PER_LANE		1	-0.0285	0.00574	24.6505	<.0001
nAVGTFACT		1	0.0108	0.00318	11.4155	0.0007
LIGHTING	Y	1	0.2856	0.1243	5.2746	0.0216
LIGHTING	P	1	1.2544	0.4622	7.3658	0.0066
Traffic_Control	3	1	0.1202	0.0503	5.7133	0.0168
Traffic_Control	2	1	-0.00315	0.037	0.0072	0.9322
Access_class	9	1	0.03	0.043	0.4862	0.4856
Access_class	7	1	-0.2467	0.0701	12.4034	0.0004
Access_class	6	1	-0.1827	0.0582	9.8575	0.0017
Access_class	5	1	-0.1295	0.0376	11.8684	0.0006
nrural_urban		1	-0.1285	0.0598	4.627	0.0315
nType_of_Shoulder	3	1	0.0853	0.0374	5.2146	0.0224
nType_of_Shoulder	2	1	-0.0338	0.039	0.7541	0.3852
Lane_width	4	1	-0.1905	0.0493	14.9527	0.0001
Lane_width	3	1	-0.2048	0.0706	8.4159	0.0037
Lane_width	2	1	-0.2104	0.0545	14.9199	0.0001
roadway_curve		1	0.2666	0.0642	17.2382	<.0001
Sidewalk_width_group	3	1	-0.235	0.0454	26.7368	<.0001
Sidewalk_width_group	2	1	-0.1609	0.0374	18.4864	<.0001
LIGHTCDE	Y	1	0.1216	0.0544	4.9992	0.0254
LIGHTCDE	P	1	-0.1976	0.051	14.9979	0.0001
Type_Friction_Course	9	1	-0.0255	0.036	0.5022	0.4786
Type_Friction_Course	5	1	-0.1846	0.079	5.4615	0.0194
Type_Friction_Course	4	1	-0.0851	0.0387	4.8203	0.0281
Type_Friction_Course	1	1	-0.3066	0.0656	21.8356	<.0001
Intersection		1	-0.1853	0.0974	3.6237	0.057
Intersect*nrural_urb		1	-0.1486	0.0556	7.1415	0.0075
Skid_Resistance		1	0.1806	0.0412	19.1908	<.0001
nrural_ur*Skid_Resis		1	-0.0845	0.0607	1.942	0.1635
Day_of_Week		1	-0.0985	0.0319	9.5222	0.002

Injury Severity Regression Model
(for a Sample of Driver Involvements of Crashes At Intersections)

Parameter		DF	Estimate	Standard Error	Wald chi-square	Pr > chi-square
Intercept		1	-1.5017	0.1458	106.0246	<.0001
Driver_Ageg_Group_x	5	1	0.5907	0.0926	40.7406	<.0001
Driver_Ageg_Group_x	4	1	0.4331	0.061	50.4666	<.0001
Driver_Ageg_Group_x	3	1	-0.281	0.0615	20.8544	<.0001
Driver_Ageg_Group_x	2	1	-0.2315	0.0661	12.2485	0.0005
Ejected_x		1	1.4373	0.1075	178.7378	<.0001
Gender_x		1	0.3613	0.0399	82.1808	<.0001
Safety_Equipment_x	1	1	-1.0658	0.0546	380.955	<.0001
Speeding_x	2	1	-0.4946	0.0628	62.0449	<.0001
Speeding_x	1	1	-0.635	0.05	161.5623	<.0001
Contributing_Cause_x	4	1	0.2591	0.0576	20.2208	<.0001
Contributing_Cause_x	3	1	0.3592	0.1852	3.7616	0.0524
Contributing_Cause_x	2	1	0.4834	0.0663	53.222	<.0001
At_Fault_driver_x		1	-0.6337	0.0562	127.3439	<.0001
Red_light_running_x		1	0.2874	0.0949	9.1735	0.0025
Residence_Code_x		1	0.2513	0.0969	6.721	0.0095
Physical_Defects_x		1	0.413	0.1159	12.7039	0.0004
Harmful_Event_Group_	7	1	0.0867	0.0611	2.0142	0.1558
Harmful_Event_Group_	6	1	0.7766	0.1069	52.7508	<.0001
Harmful_Event_Group_	5	1	-0.3227	0.1445	4.9879	0.0255
Harmful_Event_Group_	4	1	0.8009	0.0733	119.2727	<.0001
Harmful_Event_Group_	3	1	0.5945	0.0635	87.7639	<.0001
Harmful_Event_Group_	2	1	0.8169	0.1206	45.8911	<.0001
Type_of_Vehicle_x	5	1	-0.4243	0.2056	4.2575	0.0391
Type_of_Vehicle_x	4	1	0.2415	0.1179	4.1987	0.0405
Type_of_Vehicle_x	3	1	-1.1246	0.1736	41.9484	<.0001
Type_of_Vehicle_x	2	1	-0.249	0.0486	26.2746	<.0001
point_impact_x		1	0.1716	0.1104	2.4153	0.1202
point_imp*Speeding_x	2	1	0.3762	0.1793	4.4021	0.0359
point_imp*Speeding_x	1	1	0.4101	0.1268	10.4595	0.0012
Speed_limit_x	1	1	-0.4201	0.0454	85.823	<.0001
Access_class	9	1	0.00901	0.0594	0.023	0.8794
Access_class	7	1	-0.3906	0.0949	16.9464	<.0001
Access_class	6	1	-0.2795	0.0815	11.7678	0.0006

Parameter		DF	Estimate	Standard Error	Wald chi-square	Pr > chi-square
Access_class	5	1	-0.1666	0.0501	11.0488	0.0009
nrural_urban		1	-0.2953	0.0703	17.6653	<.0001
nType_of_Shoulder	3	1	0.1081	0.0508	4.529	0.0333
nType_of_Shoulder	2	1	-0.0273	0.0554	0.2425	0.6224
Lane_width	4	1	-0.2662	0.0677	15.46	<.0001
Lane_width	3	1	-0.2202	0.0924	5.6806	0.0172
Lane_width	2	1	-0.1631	0.0687	5.6431	0.0175
Sidewalk_width_group	3	1	-0.3238	0.0598	29.3069	<.0001
Sidewalk_width_group	2	1	-0.256	0.0501	26.0675	<.0001
LIGHTCDE	Y	1	0.1498	0.075	3.9869	0.0459
LIGHTCDE	P	1	-0.1358	0.068	3.988	0.0458
Type_Friction_Course	9	1	-0.0048	0.0498	0.0093	0.9232
Type_Friction_Course	5	1	-0.3204	0.1221	6.888	0.0087
Type_Friction_Course	4	1	-0.024	0.0525	0.209	0.6476
Type_Friction_Course	1	1	-0.3045	0.0907	11.2717	0.0008
Skid_Resistance		1	0.1338	0.0565	5.6056	0.0179
nrural_ur*Skid_Resis		1	-0.0311	0.0828	0.1407	0.7076

Injury Severity Regression Model
(for a Sample of Driver Involvements of Crashes at Signalized Intersections)

Parameter		DF	Estimate	Standard Error	Wald chi-square	Pr > chi-square
Intercept		1	-1.502	0.2031	54.6654	<.0001
Driver_Ageg_Group_x	5	1	0.5381	0.1344	16.0317	<.0001
Driver_Ageg_Group_x	4	1	0.3552	0.089	15.9203	<.0001
Driver_Ageg_Group_x	3	1	-0.3632	0.0884	16.8653	<.0001
Driver_Ageg_Group_x	2	1	-0.2365	0.0943	6.2886	0.0122
Gender_x		1	0.3692	0.0557	43.9095	<.0001
Safety_Equipment_x	1	1	-1.0555	0.0766	189.941	<.0001
Speeding_x	2	1	-0.323	0.0989	10.6724	0.0011
Speeding_x	1	1	-0.641	0.0713	80.8085	<.0001
At_Fault_driver_x		1	-0.5285	0.0693	58.2091	<.0001
Red_light_running_x		1	0.3642	0.1033	12.43	0.0004
Residence_Code_x		1	0.462	0.1493	9.5726	0.002
Physical_Defects_x		1	0.5547	0.153	13.137	0.0003
Ejected_x		1	1.4542	0.1621	80.4332	<.0001
Harmful_Event_Group_	7	1	-0.0184	0.0815	0.0509	0.8216
Harmful_Event_Group_	6	1	0.9469	0.1518	38.914	<.0001
Harmful_Event_Group_	5	1	-0.6168	0.2356	6.8541	0.0088
Harmful_Event_Group_	4	1	0.8346	0.1087	58.9455	<.0001
Harmful_Event_Group_	3	1	0.5	0.0908	30.3451	<.0001
Harmful_Event_Group_	2	1	0.7831	0.1637	22.8878	<.0001
point_impact_x		1	0.4213	0.082	26.4191	<.0001
Vehicle_Maneuver_x	4	1	0.0954	0.184	0.2688	0.6042
Vehicle_Maneuver_x	3	1	0.2296	0.0809	8.066	0.0045
Vehicle_Maneuver_x	2	1	-0.2741	0.0989	7.6875	0.0056
Type_of_Vehicle_x	5	1	-0.5157	0.3127	2.7207	0.0991
Type_of_Vehicle_x	4	1	0.1708	0.1812	0.8888	0.3458
Type_of_Vehicle_x	3	1	-1.0046	0.2321	18.7334	<.0001
Type_of_Vehicle_x	2	1	-0.2101	0.068	9.5585	0.002
nrural_urban		1	-0.2126	0.0564	14.1913	0.0002
Speed_limit_x	1	1	-0.3442	0.0663	26.9699	<.0001
nType_of_Shoulder	3	1	0.1503	0.0695	4.6786	0.0305
nType_of_Shoulder	2	1	-0.0471	0.0823	0.3279	0.5669
Lane_width	4	1	-0.2086	0.09	5.3735	0.0204
Lane_width	3	1	-0.3902	0.1276	9.3467	0.0022

Parameter		DF	Estimate	Standard Error	Wald chi-square	Pr > chi-square
Lane_width	**2**	1	-0.2214	0.0905	5.9819	0.0145
Access_class	**9**	1	0.039	0.0851	0.2103	0.6465
Access_class	**7**	1	-0.4751	0.138	11.8589	0.0006
Access_class	**6**	1	-0.4326	0.121	12.7897	0.0003
Access_class	**5**	1	-0.1797	0.0682	6.9412	0.0084
Type_Friction_Course	**9**	1	-0.0879	0.0708	1.5392	0.2147
Type_Friction_Course	**5**	1	-0.6178	0.2174	8.0718	0.0045
Type_Friction_Course	**4**	1	-0.0117	0.0734	0.0255	0.8731
Type_Friction_Course	**1**	1	-0.3755	0.1368	7.531	0.0061
Sidewalk_width_group	**3**	1	-0.4082	0.081	25.376	<.0001
Sidewalk_width_group	**2**	1	-0.362	0.0684	27.9692	<.0001

Injury Severity Regression Model
(for a Sample of Driver Involvements of Crashes in Road Segments and Unsignalized Intersections)

Parameter		DF	Estimate	Standard Error	Wald chi-square	Pr > chi-square
Intercept		1	-1.5058	0.1811	69.0951	<.0001
Driver_Ageg_Group_x	5	1	0.4104	0.0973	17.7749	<.0001
Driver_Ageg_Group_x	4	1	0.3434	0.0584	34.586	<.0001
Driver_Ageg_Group_x	3	1	-0.2235	0.0494	20.4331	<.0001
Driver_Ageg_Group_x	2	1	-0.2851	0.0557	26.2288	<.0001
Ejected_x		1	1.4773	0.0783	355.5834	<.0001
Speeding_x	2	1	-0.1925	0.124	2.4089	0.1206
Speeding_x	1	1	-1.0001	0.0958	108.9808	<.0001
Gender_x		1	0.1894	0.0738	6.58	0.0103
Safety_Equipment_x	1	1	-1.1979	0.0546	481.8771	<.0001
Gender_x*Safety_Equi	1	1	0.2263	0.0833	7.3866	0.0066
At_Fault_driver_x		1	-0.6592	0.0458	207.2838	<.0001
Physical_Defects_x		1	0.4034	0.0873	21.3638	<.0001
Harmful_Event_Group_	7	1	0.1242	0.0525	5.5961	0.018
Harmful_Event_Group_	6	1	0.6095	0.0735	68.8627	<.0001
Harmful_Event_Group_	5	1	-0.1787	0.1006	3.1536	0.0758
Harmful_Event_Group_	4	1	0.706	0.0788	80.341	<.0001
Harmful_Event_Group_	3	1	0.5824	0.0591	97.0609	<.0001
Harmful_Event_Group_	2	1	1.1667	0.1007	134.2465	<.0001
Contributing_Cause_x	4	1	0.6373	0.0462	190.6432	<.0001
Contributing_Cause_x	3	1	0.8008	0.1243	41.4915	<.0001
Contributing_Cause_x	2	1	0.58	0.0578	100.6664	<.0001
Type_of_Vehicle_x	5	1	-0.2254	0.1608	1.9665	0.1608
Type_of_Vehicle_x	4	1	-0.0359	0.0883	0.1648	0.6848
Type_of_Vehicle_x	3	1	-1.015	0.1636	38.4739	<.0001
Type_of_Vehicle_x	2	1	-0.1861	0.042	19.6509	<.0001
point_impact_x		1	0.0894	0.0818	1.1951	0.2743
point_imp*Speeding_x	2	1	0.0886	0.1632	0.2945	0.5873
point_imp*Speeding_x	1	1	0.2806	0.1022	7.5403	0.006
Off_Roadway		1	-0.3029	0.0679	19.8994	<.0001
Off_Roadw*Speeding_x	2	1	-0.1102	0.1335	0.6809	0.4093
Off_Roadw*Speeding_x	1	1	0.4169	0.1027	16.4861	<.0001
nWork_Area_x	3	1	-0.1873	0.119	2.4794	0.1153
nWork_Area_x	2	1	-0.298	0.1083	7.5735	0.0059
Private_vehicle_use_		1	0.448	0.133	11.3421	0.0008

Parameter		DF	Estimate	Standard Error	Wald chi-square	Pr > chi-square
Private_ve*nAVGTFACT		1	-0.0256	0.0102	6.3133	0.012
Speed_limit_x	1	1	-0.4195	0.0413	103.1739	<.0001
ADT_PER_LANE		1	-0.0411	0.00684	36.0999	<.0001
nAVGTFACT		1	0.0338	0.00981	11.8843	0.0006
LIGHTING	Y	1	0.3385	0.1538	4.8422	0.0278
LIGHTING	P	1	1.311	0.6635	3.9039	0.0482
Traffic_Control	3	1	0.2115	0.0485	19.0094	<.0001
Traffic_Control	2	1	0.2769	0.182	2.314	0.1282
nrural_urban		1	-0.2033	0.0635	10.2391	0.0014
Lane_width	4	1	-0.1706	0.0603	8.0143	0.0046
Lane_width	3	1	-0.1349	0.0872	2.3916	0.122
Lane_width	2	1	-0.2012	0.0702	8.2139	0.0042
roadway_curve		1	0.33	0.0681	23.5057	<.0001
Sidewalk_width_group	3	1	-0.2224	0.0531	17.5614	<.0001
Sidewalk_width_group	2	1	-0.117	0.0424	7.6027	0.0058
LIGHTCDE	Y	1	0.114	0.0685	2.7732	0.0959
LIGHTCDE	P	1	-0.1784	0.062	8.2748	0.004
Type_Friction_Course	9	1	-0.0081	0.0427	0.036	0.8495
Type_Friction_Course	5	1	-0.1017	0.0862	1.3923	0.238
Type_Friction_Course	4	1	-0.0736	0.0437	2.8368	0.0921
Type_Friction_Course	1	1	-0.2825	0.0756	13.9796	0.0002
Skid_Resistance		1	0.2283	0.0486	22.0224	<.0001
nrural_ur*Skid_Resis		1	-0.141	0.074	3.6325	0.0567
AUX_Lane_Num	3	1	0.072	0.0611	1.3875	0.2388
AUX_Lane_Num	2	1	-0.1258	0.052	5.8525	0.0156
AUX_Lane_Num	1	1	0.0528	0.0377	1.9602	0.1615

Injury Severity Regression Model
(for a Sample of Driver Involvements of Crashes in Road Segments)

Parameter		DF	Estimate	Standard Error	Wald chi-square	Pr > chi-square
Intercept		1	-1.1598	0.1538	56.8505	<.0001
Driver_Ageg_Group_x	5	1	0.1439	0.1551	0.8613	0.3534
Driver_Ageg_Group_x	4	1	0.2255	0.0818	7.6	0.0058
Driver_Ageg_Group_x	3	1	-0.22	0.0604	13.2791	0.0003
Driver_Ageg_Group_x	2	1	-0.288	0.0696	17.1175	<.0001
Ejected_x		1	1.5113	0.0939	258.9663	<.0001
Speeding_x	2	1	-0.1444	0.1366	1.118	0.2904
Speeding_x	1	1	-0.9115	0.1025	79.1546	<.0001
Gender_x		1	0.3764	0.0455	68.5432	<.0001
Safety_Equipment_x	1	1	-1.1328	0.0558	411.3992	<.0001
At_Fault_driver_x		1	-0.6357	0.0566	126.3422	<.0001
Physical_Defects_x		1	0.4257	0.1009	17.8134	<.0001
Harmful_Event_Group_	7	1	0.1386	0.0632	4.8061	0.0284
Harmful_Event_Group_	6	1	0.6308	0.0832	57.4984	<.0001
Harmful_Event_Group_	5	1	-0.1008	0.1178	0.7315	0.3924
Harmful_Event_Group_	4	1	0.8354	0.1279	42.6461	<.0001
Harmful_Event_Group_	3	1	0.5978	0.0786	57.7835	<.0001
Harmful_Event_Group_	2	1	1.3997	0.1226	130.3463	<.0001
Contributing_Cause_x	4	1	0.6973	0.0561	154.2651	<.0001
Contributing_Cause_x	3	1	0.8338	0.1436	33.7086	<.0001
Contributing_Cause_x	2	1	0.4975	0.0792	39.4087	<.0001
Type_of_Vehicle_x	5	1	-0.1937	0.1953	0.9829	0.3215
Type_of_Vehicle_x	4	1	-0.0774	0.1055	0.5388	0.4629
Type_of_Vehicle_x	3	1	-0.9904	0.1567	39.9449	<.0001
Type_of_Vehicle_x	2	1	-0.1369	0.0526	6.7632	0.0093
Off_Roadway		1	-0.3344	0.0753	19.6953	<.0001
Off_Roadw*Speeding_x	2	1	-0.0575	0.1524	0.1422	0.7061
Off_Roadw*Speeding_x	1	1	0.3276	0.1147	8.1547	0.0043
Speed_limit_x	1	1	-0.3898	0.0556	49.149	<.0001
ADT_PER_LANE		1	-0.0534	0.00831	41.26	<.0001
nAVGTFACT		1	0.0119	0.00415	8.2304	0.0041
nrural_urban		1	-0.0879	0.0792	1.2313	0.2672
roadway_curve		1	0.292	0.0816	12.7964	0.0003
Sidewalk_width_group	3	1	-0.1854	0.0685	7.3244	0.0068
Sidewalk_width_group	2	1	-0.0939	0.054	3.0214	0.0822

Parameter		DF	Estimate	Standard Error	Wald chi-square	Pr > chi-square
LIGHTCDE	**Y**	1	0.0515	0.0847	0.3699	0.5431
LIGHTCDE	**P**	1	-0.2613	0.0797	10.7481	0.001
Type_Friction_Course	**9**	1	-0.0535	0.0539	0.9831	0.3214
Type_Friction_Course	**5**	1	-0.0676	0.1058	0.409	0.5225
Type_Friction_Course	**4**	1	-0.1184	0.0562	4.4327	0.0353
Type_Friction_Course	**1**	1	-0.3248	0.0971	11.1878	0.0008
Skid_Resistance		1	0.2659	0.0617	18.5609	<.0001
nrural_ur*Skid_Resis		1	-0.1925	0.093	4.2805	0.0386

Injury Severity Regression Model
(for Involvements of Crashes at Unsignalized Intersections)

Parameter		DF	Estimate	Standard Error	Wald chi-square	Pr > ch-square
Intercept		1	-1.1806	0.1517	60.6029	<.0001
Driver_Ageg_Group_x	5	1	0.6533	0.1276	26.1955	<.0001
Driver_Ageg_Group_x	4	1	0.5099	0.0837	37.0918	<.0001
Driver_Ageg_Group_x	3	1	-0.1945	0.0862	5.0914	0.024
Driver_Ageg_Group_x	2	1	-0.2239	0.0927	5.8371	0.0157
Gender_x		1	0.3572	0.0571	39.0934	<.0001
Safety_Equipment_x	1	1	-1.0609	0.0778	185.7502	<.0001
Speeding_x	2	1	-0.4986	0.0894	31.1132	<.0001
Speeding_x	1	1	-0.5562	0.0653	72.5139	<.0001
Contributing_Cause_x	4	1	0.433	0.0826	27.478	<.0001
Contributing_Cause_x	3	1	0.6865	0.2544	7.2837	0.007
Contributing_Cause_x	2	1	0.6197	0.0896	47.8621	<.0001
At_Fault_driver_x		1	-0.6733	0.0786	73.3714	<.0001
Ejected_x		1	1.4279	0.1447	97.3116	<.0001
Harmful_Event_Group_	7	1	0.0961	0.0933	1.0594	0.3033
Harmful_Event_Group_	6	1	0.4924	0.1548	10.1252	0.0015
Harmful_Event_Group_	5	1	-0.2104	0.1891	1.2378	0.2659
Harmful_Event_Group_	4	1	0.591	0.1113	28.1834	<.0001
Harmful_Event_Group_	3	1	0.5621	0.0932	36.3935	<.0001
Harmful_Event_Group_	2	1	0.7264	0.1799	16.3001	<.0001
point_impact_x		1	0.467	0.0701	44.4169	<.0001
Type_of_Vehicle_x	5	1	-0.3453	0.2724	1.6073	0.2049
Type_of_Vehicle_x	4	1	0.3019	0.1568	3.7102	0.0541
Type_of_Vehicle_x	3	1	-1.2987	0.2609	24.7697	<.0001
Type_of_Vehicle_x	2	1	-0.2987	0.0697	18.3765	<.0001
nrural_urban		1	-0.4328	0.0579	55.8018	<.0001
roadway_curve		1	0.4124	0.1221	11.3989	0.0007
Speed_limit_x	1	1	-0.4744	0.0616	59.4099	<.0001
ADT_PER_LANE		1	-0.0288	0.0104	7.7464	0.0054
AUX_Lane_Num	3	1	0.174	0.1029	2.8599	0.0908
AUX_Lane_Num	2	1	-0.1512	0.0877	2.9696	0.0848
AUX_Lane_Num	1	1	0.1273	0.0611	4.335	0.0373
Sidewalk_width_group	3	1	-0.2717	0.0808	11.3154	0.0008
Sidewalk_width_group	2	1	-0.1823	0.0653	7.8025	0.0052

Odds Ratio Estimates
(for the Six Road Entity Models)

Variable	Level	Overall	Inters	Signal	Segment	Pure Segment	Non-signal
Driver_Ageg_Group_x	5	1.621	1.805	1.713	1.507	1.155	1.922
Driver_Ageg_Group_x	4	1.422	1.542	1.426	1.410	1.253	1.665
Driver_Ageg_Group_x	3	0.779	0.755	0.695	0.800	0.803	0.823
Driver_Ageg_Group_x	2	0.767	0.793	0.789	0.752	0.750	0.799
Ejected_x		4.270	4.209	4.281	4.381	4.533	4.170
Gender_x		1.217	1.435	1.447	1.209	1.457	1.429
Safety_Equipment_x	1	0.303	0.344	0.348	0.302	0.322	0.346
Gender_x*Safety_Equi	1	1.245			1.254		
Speeding_x	2	0.863	0.610	0.724	0.825	0.866	0.607
Speeding_x	1	0.409	0.530	0.527	0.368	0.402	0.573
Contributing_Cause_x	4	1.605	1.296		1.891	2.008	1.542
Contributing_Cause_x	3	1.748	1.432		2.227	2.302	1.987
Contributing_Cause_x	2	1.593	1.622		1.786	1.645	1.858
At_Fault_driver_x		0.538	0.531	0.589	0.517	0.530	0.510
Red_light_running_x			1.333	1.439			
Residence_Code_x		1.175	1.286	1.587			
Physical_Defects_x		1.535	1.511	1.741	1.497	1.531	
Harmful_Event_Group_	7	1.097	1.091	0.982	1.132	1.149	1.101
Harmful_Event_Group_	6	1.810	2.174	2.578	1.840	1.879	1.636
Harmful_Event_Group_	5	0.779	0.724	0.540	0.836	0.904	0.810
Harmful_Event_Group_	4	2.242	2.228	2.304	2.026	2.306	1.806
Harmful_Event_Group_	3	1.784	1.812	1.649	1.790	1.818	1.754
Harmful_Event_Group_	2	2.875	2.263	2.188	3.211	4.054	2.068
Vehicle_Maneuver_x	4			1.100			
Vehicle_Maneuver_x	3			1.258			
Vehicle_Maneuver_x	2			0.760			
Type_of_Vehicle_x	5	0.756	0.654	0.597	0.798	0.824	0.708
Type_of_Vehicle_x	4	1.050	1.273	1.186	0.965	0.926	1.352
Type_of_Vehicle_x	3	0.357	0.325	0.366	0.362	0.371	0.273
Type_of_Vehicle_x	2	0.820	0.780	0.811	0.830	0.872	0.742
point_impact_x		1.091	1.187	1.524	1.094		1.595
point_imp*Speeding_x	2	1.240	1.457		1.093		
point_imp*Speeding_x	1	1.412	1.507		.		
Off_Roadway		0.613			0.739	0.716	
Off_Roadw*Speeding_x	2	0.783			0.896	0.944	
Off_Roadw*Speeding_x	1	1.289			1.517	1.388	

Variable	Level	Overall	Inters	Signal	Segment	Pure Segment	Non-signal
Off_Roadw*Multivehic	1	2.033					
nWork_Area_x	3	0.826			0.829		
nWork_Area_x	2	0.750			0.742		
Private_vehicle_use_					1.565		
Private_ve*nAVGTFACT					0.975		
Multivehicle	1	0.469					
Intersect*Multivehic	1	1.476					
Speed_limit_x	1	0.676	0.657	0.709	0.657	0.677	0.622
ADT_PER_LANE		0.972			0.960	0.948	0.972
nAVGTFACT		1.011			1.034	1.012	
LIGHTING	Y	1.331			1.403		
LIGHTING	P	3.506			3.710		
Traffic_Control	3	1.128			1.236		
Traffic_Control	2	0.997			1.319		
Access_class	9	1.030	1.009	1.040			
Access_class	7	0.781	0.677	0.622			
Access_class	6	0.833	0.756	0.649			
Access_class	5	0.879	0.847	0.836			
nrural_urban		0.879	0.744	0.808	0.816	0.916	0.649
nType_of_Shoulder	3	1.089	1.114	1.162			
nType_of_Shoulder	2	0.967	0.973	0.954			
Lane_width	4	0.827	0.766	0.812	0.843		
Lane_width	3	0.815	0.802	0.677	0.874		
Lane_width	2	0.810	0.850	0.801	0.818		
roadway_curve		1.306			1.391	1.339	1.510
Sidewalk_width_group	3	0.791	0.723	0.665	0.801	0.831	0.762
Sidewalk_width_group	2	0.851	0.774	0.696	0.890	0.910	0.833
LIGHTCDE	Y	1.129	1.162		1.121	1.053	
LIGHTCDE	P	0.821	0.873		0.837	0.770	
Type_Friction_Course	9	0.975	0.995	0.916	0.992	0.948	
Type_Friction_Course	5	0.831	0.726	0.539	0.903	0.935	
Type_Friction_Course	4	0.918	0.976	0.988	0.929	0.888	
Type_Friction_Course	1	0.736	0.737	0.687	0.754	0.723	
Intersection		0.831					
Intersect*nrural_urb		0.862					
Skid_Resistance		1.198	1.143		1.256	1.305	
nrural_ur*Skid_Resis		0.919	0.969		0.868	0.825	
AUX_Lane_Num	3				1.075		1.190

Variable	Level	Overall	Inters	Signal	Segment	Pure Segment	Non-signal
AUX_Lane_Num	**2**				0.882		0.860
AUX_Lane_Num	**1**				1.054		1.136
Day_of_Week		0.906					

APPENDIX E: INJURY SEVERITY MODELS' FINAL ANALYSIS – CRASH TYPES MODELS

Injury Severity Regression Model
(for Driver Involvements in All Rear-end Crashes)

Parameter		DF	Estimate	Standard Error	Wald chi-square	Pr > chi-square
Intercept		1	-0.5609	0.2769	4.1035	0.0428
Driver_Ageg_Group_x	**5**	1	0.0581	0.2228	0.0679	0.7944
Driver_Ageg_Group_x	**4**	1	0.2053	0.1141	3.2355	0.0721
Driver_Ageg_Group_x	**3**	1	-0.4641	0.1025	20.5118	<.0001
Driver_Ageg_Group_x	**2**	1	-0.4699	0.1148	16.7433	<.0001
Gender_x		1	0.0618	0.1582	0.1526	0.696
Safety_Equipment_x	**1**	1	-1.3669	0.1075	161.5832	<.0001
Speeding_x	**2**	1	-0.4582	0.0906	25.5751	<.0001
Speeding_x	**1**	1	-0.8950	0.0804	123.9842	<.0001
At_Fault_driver_x		1	-0.5293	0.0743	50.7892	<.0001
Residence_Code_x		1	0.5711	0.1947	8.6043	0.0034
Physical_Defects_x		1	0.7787	0.1619	23.1391	<.0001
Ejected_x		1	0.5970	0.2415	6.1132	0.0134
Gender_x*Safety_Equi	**1**	1	0.4635	0.1720	7.2598	0.0071
Type_of_Vehicle_x	**5**	1	-0.0584	0.3503	0.0278	0.8676
Type_of_Vehicle_x	**4**	1	0.8410	0.2455	11.7381	0.0006
Type_of_Vehicle_x	**3**	1	-1.2242	0.2871	18.1846	<.0001
Type_of_Vehicle_x	**2**	1	-0.1810	0.0753	5.7833	0.0162
nrural_urban		1	-0.2087	0.0672	9.6569	0.0019
nWork_Area_x	**3**	1	-0.2607	0.2397	1.1825	0.2768
nWork_Area_x	**2**	1	-0.6349	0.2225	8.1401	0.0043
Concrete_Surface		1	-0.7592	0.3272	5.3832	0.0203
Speed_limit_x	**1**	1	-0.5212	0.0817	40.6879	<.0001
ADT_PER_LANE		1	-0.0548	0.0126	19.0552	<.0001
Skid_Resistance		1	0.2713	0.0703	14.9093	0.0001
Lane_width	**4**	1	-0.2847	0.1128	6.3714	0.0116
Lane_width	**3**	1	-0.4168	0.1810	5.3007	0.0213
Lane_width	**2**	1	-0.3893	0.1287	9.1579	0.0025
LIGHTCDE	**Y**	1	0.3065	0.1059	8.3814	0.0038
LIGHTCDE	**P**	1	-0.2811	0.1128	6.2051	0.0127
LIGHTING	**Y**	1	0.2673	0.2259	1.4004	0.2367
LIGHTING	**P**	1	1.3902	0.5775	5.7949	0.0161
Access_class	**9**	1	0.0495	0.0963	0.2636	0.6077
Access_class	**7**	1	-0.2655	0.1703	2.4309	0.119

Parameter		DF	Estimate	Standard Error	Wald chi-square	Pr > chi-square
Access_class	**6**	1	-0.2927	0.1389	4.4397	0.0351
Access_class	**5**	1	-0.1622	0.0811	4.0012	0.0455
Type_Friction_Course	**9**	1	0.0364	0.0795	0.2093	0.6473
Type_Friction_Course	**5**	1	-0.2273	0.1865	1.4860	0.2228
Type_Friction_Course	**4**	1	-0.2150	0.0898	5.7367	0.0166
Type_Friction_Course	**1**	1	-0.3308	0.1519	4.7430	0.0294
Sidewalk_width_group	**3**	1	-0.2559	0.0978	6.8516	0.0089
Sidewalk_width_group	**2**	1	-0.1896	0.0775	5.9773	0.0145
Day_of_Week		1	-0.2748	0.0798	11.8693	0.0006

Injury Severity Regression Model
(for Driver Involvements in Rear-end Crashes at Signalized Intersections)

Parameter		DF	Estimate	Standard Error	Wald chi-square	Pr > chi-square
Intercept		1	-1.2322	0.4810	6.5637	0.0104
driver_ageg_group_x	4	1	0.2493	0.1836	1.8426	0.1746
driver_ageg_group_x	3	1	-0.3831	0.1969	3.7870	0.0517
driver_ageg_group_x	2	1	-0.4931	0.2351	4.3997	0.0359
gender_x		1	-0.4110	0.1171	12.3059	0.0005
Safety_Equipment_x	1	1	-1.2121	0.1480	67.0601	<.0001
Speeding_x	2	1	-0.5367	0.1767	9.2249	0.0024
Speeding_x	1	1	-1.1693	0.1619	52.1472	<.0001
At_Fault_driver_x		1	-0.5669	0.1461	15.0499	0.0001
Residence_Code_x		1	0.9107	0.4177	4.7539	0.0292
Physical_Defects_x		1	0.8145	0.2718	8.9801	0.0027
Ejected_x		1	1.4877	0.3679	16.3557	<.0001
Speed_limit_x	1	1	-0.5289	0.1411	14.0420	0.0002
Lane_width	4	1	-0.3941	0.2031	3.7674	0.0523
Lane_width	3	1	-0.7658	0.3690	4.3075	0.0379
Lane_width	2	1	-0.3597	0.2160	2.7722	0.0959
Sidewalk_width_group	3	1	-0.4098	0.1654	6.1411	0.0132
Sidewalk_width_group	2	1	-0.5384	0.1348	15.9391	<.0001

Injury Severity Regression Model
(for Driver Involvements in Rear-end Crashes at Unsignalized Intersections)

Parameter		DF	Estimate	Standard Error	Wald chi-square	Pr > chi-square
Intercept		1	-0.2571	0.3404	0.5705	0.4501
Driver_Ageg_Group_x	5	1	-0.3890	0.6039	0.4150	0.5194
Driver_Ageg_Group_x	4	1	0.4891	0.2412	4.1121	0.0426
Driver_Ageg_Group_x	3	1	-0.8625	0.2814	9.3932	0.0022
Driver_Ageg_Group_x	2	1	-0.3653	0.2496	2.1410	0.1434
Gender_x		1	-0.3844	0.3883	0.9799	0.3222
Safety_Equipment_x	1	1	-1.7666	0.2542	48.3044	<.0001
Speeding_x	2	1	-0.2897	0.2012	2.0734	0.1499
Speeding_x	1	1	-0.6692	0.1834	13.3127	0.0003
Gender_x*Safety_Equi	1	1	1.1815	0.4221	7.8361	0.0051
point_impact_x		1	1.0508	0.4958	4.4909	0.0341
Type_of_Vehicle_x	5	1	0.1979	0.7460	0.0704	0.7907
Type_of_Vehicle_x	4	1	1.0859	0.3747	8.3978	0.0038
Type_of_Vehicle_x	3	1	-1.6438	0.7451	4.8667	0.0274
Type_of_Vehicle_x	2	1	-0.3251	0.1820	3.1914	0.074
nrural_urban		1	-0.5334	0.1577	11.4404	0.0007
Speed_limit_x	1	1	-0.7741	0.1666	21.5910	<.0001
ADT_PER_LANE		1	-0.0614	0.0294	4.3714	0.0365
LIGHTCDE	Y	1	0.6465	0.2384	7.3517	0.0067
LIGHTCDE	P	1	-0.1629	0.2673	0.3715	0.5422

Injury Severity Regression Model
(for Driver Involvements in Rear-end Crashes at Unsignalized Intersections and Road Segments)

Parameter		DF	Estimate	Standard Error	Wald chi-square	Pr > chi-square
Intercept		1	-0.0586	0.2851	0.0423	0.8371
Driver_Ageg_Group_x	5	1	-0.1449	0.3322	0.1903	0.6627
Driver_Ageg_Group_x	4	1	0.0809	0.1539	0.2761	0.5993
Driver_Ageg_Group_x	3	1	-0.5358	0.1295	17.1241	<.0001
Driver_Ageg_Group_x	2	1	-0.4619	0.1411	10.7245	0.0011
Gender_x		1	0.00229	0.2060	0.0001	0.9911
Safety_Equipment_x	1	1	-1.4842	0.1382	115.2678	<.0001
Speeding_x	2	1	-0.3941	0.1154	11.6691	0.0006
Speeding_x	1	1	-0.7772	0.0988	61.8782	<.0001
At_Fault_driver_x		1	-0.5197	0.0929	31.2662	<.0001
Physical_Defects_x		1	0.7982	0.2214	13.0041	0.0003
Gender_x*Safety_Equi	1	1	0.6460	0.2238	8.3343	0.0039
Type_of_Vehicle_x	5	1	-0.1438	0.4673	0.0948	0.7582
Type_of_Vehicle_x	4	1	1.2934	0.2032	40.5131	<.0001
Type_of_Vehicle_x	3	1	-1.2014	0.3354	12.8335	0.0003
Type_of_Vehicle_x	2	1	-0.1621	0.0969	2.8026	0.0941
nrural_urban		1	-0.2350	0.0862	7.4361	0.0064
nWork_Area_x	3	1	-0.1121	0.2744	0.1670	0.6828
nWork_Area_x	2	1	-0.7976	0.3132	6.4875	0.0109
Concrete_Surface		1	-0.8725	0.4263	4.1899	0.0407
Speed_limit_x	1	1	-0.5945	0.1028	33.4161	<.0001
ADT_PER_LANE		1	-0.0738	0.0160	21.1855	<.0001
nAVGTFACT		1	0.0189	0.00962	3.8729	0.0491
Skid_Resistance		1	0.3424	0.0901	14.4380	0.0001
LIGHTCDE	Y	1	0.3749	0.1342	7.8109	0.0052
LIGHTCDE	P	1	-0.2934	0.1460	4.0366	0.0445
LIGHTING	Y	1	0.4732	0.2859	2.7401	0.0979
LIGHTING	P	1	1.5979	0.8186	3.8106	0.0509
Type_Friction_Course	9	1	0.0578	0.1002	0.3320	0.5645
Type_Friction_Course	5	1	-0.0122	0.2101	0.0034	0.9536
Type_Friction_Course	4	1	-0.2694	0.1112	5.8701	0.0154
Type_Friction_Course	1	1	-0.3751	0.1838	4.1645	0.0413
Sidewalk_width_group	3	1	-0.3608	0.1276	7.9937	0.0047
Sidewalk_width_group	2	1	-0.1504	0.0992	2.2978	0.1296

Parameter		DF	Estimate	Standard Error	Wald chi-square	Pr > chi-square
Day_of_Week		1	-0.3330	0.1064	9.7889	0.0018

Injury Severity Regression Model
(for Driver Involvements in Rear-end Crashes on Road Segments)

Parameter		DF	Estimate	Standard Error	Wald chi-square	Pr > chi-square
Intercept		1	-1.1824	0.4549	6.7575	0.0093
Driver_Ageg_Group_x	5	1	-0.0901	0.3984	0.0512	0.821
Driver_Ageg_Group_x	4	1	-0.1790	0.2026	0.7811	0.3768
Driver_Ageg_Group_x	3	1	-0.4782	0.1467	10.6235	0.0011
Driver_Ageg_Group_x	2	1	-0.5345	0.1706	9.8162	0.0017
Gender_x		1	0.4987	0.1010	24.3724	<.0001
Safety_Equipment_x	1	1	-1.1835	0.1321	80.2867	<.0001
Speeding_x	2	1	-0.3882	0.1365	8.0891	0.0045
Speeding_x	1	1	-0.8530	0.1174	52.7805	<.0001
At_Fault_driver_x		1	-0.6034	0.1113	29.3683	<.0001
Physical_Defects_x		1	0.8725	0.2534	11.8543	0.0006
Type_of_Vehicle_x	5	1	-0.0169	0.6027	0.0008	0.9776
Type_of_Vehicle_x	4	1	1.5337	0.2351	42.5751	<.0001
Type_of_Vehicle_x	3	1	-0.6172	0.4428	1.9428	0.1634
Type_of_Vehicle_x	2	1	-0.0950	0.1148	0.6855	0.4077
Private_vehicle_use_		1	0.7025	0.3334	4.4394	0.0351
Concrete_Surface		1	-1.5191	0.7148	4.5168	0.0336
Speed_limit_x	1	1	-0.6768	0.1129	35.9072	<.0001
ADT_PER_LANE		1	-0.0769	0.0189	16.5112	<.0001
nAVGTFACT		1	0.0312	0.0105	8.7900	0.003
Skid_Resistance		1	0.4331	0.1069	16.4154	<.0001
LIGHTCDE	Y	1	0.2822	0.1593	3.1393	0.0764
LIGHTCDE	P	1	-0.3241	0.1714	3.5755	0.0586
Type_Friction_Course	9	1	0.0904	0.1177	0.5904	0.4423
Type_Friction_Course	5	1	0.0621	0.2376	0.0682	0.7939
Type_Friction_Course	4	1	-0.3725	0.1346	7.6633	0.0056
Type_Friction_Course	1	1	-0.5271	0.2363	4.9764	0.0257
Day_of_Week		1	-0.3308	0.1263	6.8613	0.0088

Injury Severity Regression Model
(for Driver Involvements in All Angle Crashes)

Parameter		DF	Estimate	Standard Error	Wald chi-square	Pr > chi-square
Intercept		1	-1.2181	0.2705	20.2723	<.0001
Driver_Ageg_Group_x	5	1	0.7416	0.1282	33.4500	<.0001
Driver_Ageg_Group_x	4	1	0.5088	0.0915	30.9554	<.0001
Driver_Ageg_Group_x	3	1	-0.1601	0.0975	2.6946	0.1007
Driver_Ageg_Group_x	2	1	-0.0588	0.1004	0.3430	0.5581
Gender_x		1	0.3691	0.0630	34.3241	<.0001
Safety_Equipment_x	1	1	-1.2004	0.0849	200.0428	<.0001
Speeding_x	2	1	-0.3671	0.1162	9.9712	0.0016
Speeding_x	1	1	-0.4457	0.0714	39.0068	<.0001
Contributing_Cause_x	4	1	0.3932	0.0944	17.3388	<.0001
Contributing_Cause_x	3	1	0.9424	0.3109	9.1864	0.0024
Contributing_Cause_x	2	1	0.5754	0.0974	34.9059	<.0001
At_Fault_driver_x		1	-0.7578	0.0851	79.3744	<.0001
Red_light_running_x		1	0.2818	0.1291	4.7622	0.0291
Physical_Defects_x		1	0.4373	0.1894	5.3301	0.021
Ejected_x		1	1.4262	0.1727	68.2152	<.0001
Off_Roadway		1	0.4553	0.2266	4.0393	0.0445
point_impact_x		1	0.5828	0.0695	70.3902	<.0001
Type_of_Vehicle_x	5	1	-0.3793	0.3316	1.3085	0.2527
Type_of_Vehicle_x	4	1	0.0925	0.1934	0.2287	0.6325
Type_of_Vehicle_x	3	1	-1.3706	0.2995	20.9486	<.0001
Type_of_Vehicle_x	2	1	-0.2320	0.0774	8.9798	0.0027
nrural_urban		1	-0.3729	0.0643	33.5794	<.0001
Intersection		1	0.2742	0.0711	14.8783	0.0001
Speed_limit_x	1	1	-0.3817	0.0680	31.5207	<.0001
ADT_PER_LANE		1	-0.0320	0.0119	7.2040	0.0073
Lane_width	4	1	-0.3058	0.1155	7.0128	0.0081
Lane_width	3	1	-0.1802	0.1464	1.5157	0.2183
Lane_width	2	1	-0.1961	0.1155	2.8834	0.0895
Shoulder_width	4	1	-0.1775	0.1570	1.2777	0.2583
Shoulder_width	3	1	-0.4323	0.2087	4.2921	0.0383
Shoulder_width	2	1	-0.4580	0.2083	4.8349	0.0279
Access_class	9	1	-0.0989	0.0870	1.2922	0.2556
Access_class	7	1	-0.4734	0.1416	11.1782	0.0008
Access_class	6	1	-0.3644	0.1191	9.3670	0.0022

Parameter		DF	Estimate	Standard Error	Wald chi-square	Pr > chi-square
Access_class	**5**	1	-0.2799	0.0777	12.9794	0.0003

Injury Severity Regression Model
(for Driver Involvements in Angle Crashes at Signalized Intersections)

Parameter		DF	Estimate	Standard Error	Wald chi-square	Pr > chi-square
Intercept		1	-0.9624	0.2321	17.2015	<.0001
Driver_Ageg_Group_x	**5**	1	0.9097	0.2269	16.0807	<.0001
Driver_Ageg_Group_x	**4**	1	0.3669	0.1724	4.5291	0.0333
Driver_Ageg_Group_x	**3**	1	-0.4071	0.1853	4.8273	0.028
Driver_Ageg_Group_x	**2**	1	-0.0171	0.1763	0.0094	0.9227
Gender_x		1	0.5176	0.1093	22.4411	<.0001
Safety_Equipment_x	**1**	1	-1.1770	0.1510	60.7355	<.0001
Speeding_x	**2**	1	-0.2305	0.1986	1.3470	0.2458
Speeding_x	**1**	1	-0.3713	0.1188	9.7593	0.0018
Ejected_x		1	1.6944	0.2564	43.6869	<.0001
point_impact_x		1	0.6300	0.1320	22.7801	<.0001
nrural_urban		1	-0.3107	0.1147	7.3375	0.0068
Speed_limit_x	**1**	1	-0.4499	0.1219	13.6209	0.0002
nType_of_Shoulder	**3**	1	0.5071	0.1468	11.9269	0.0006
nType_of_Shoulder	**2**	1	-0.0523	0.1783	0.0860	0.7694
Access_class	**9**	1	-0.1286	0.1591	0.6529	0.4191
Access_class	**7**	1	-0.5482	0.2402	5.2083	0.0225
Access_class	**6**	1	-1.0192	0.2721	14.0283	0.0002
Access_class	**5**	1	-0.3595	0.1398	6.6163	0.0101
Sidewalk_width_group	**3**	1	-0.5469	0.1668	10.7518	0.001
Sidewalk_width_group	**2**	1	-0.3290	0.1382	5.6624	0.0173

Injury Severity Regression Model
(for Driver Involvements in Angle Crashes at Unsignalized Intersections)

Parameter		DF	Estimate	Standard Error	Wald chi-square	Pr > chi-square
Intercept		1	-1.8080	0.3197	31.9845	<.0001
Driver_Ageg_Group_x	5	1	0.9699	0.1837	27.8763	<.0001
Driver_Ageg_Group_x	4	1	0.6886	0.1354	25.8591	<.0001
Driver_Ageg_Group_x	3	1	0.0559	0.1504	0.1379	0.7103
Driver_Ageg_Group_x	2	1	0.0335	0.1538	0.0473	0.8278
Gender_x		1	0.4217	0.0953	19.5616	<.0001
Safety_Equipment_x	1	1	-1.1885	0.1284	85.6882	<.0001
Speeding_x	2	1	-0.7903	0.2034	15.0902	0.0001
Speeding_x	1	1	-0.4986	0.1183	17.7655	<.0001
Contributing_Cause_x	4	1	0.5735	0.1588	13.0474	0.0003
Contributing_Cause_x	3	1	1.7850	0.4678	14.5627	0.0001
Contributing_Cause_x	2	1	0.8968	0.1497	35.9086	<.0001
At_Fault_driver_x		1	-1.1702	0.1290	82.3447	<.0001
Ejected_x		1	1.6699	0.2129	61.5207	<.0001
point_impact_x		1	0.7315	0.1046	48.8672	<.0001
Private_vehicle_use_		1	0.7028	0.2748	6.5395	0.0106
nrural_urban		1	-0.5616	0.0982	32.7103	<.0001
Speed_limit_x	1	1	-0.4990	0.0984	25.7142	<.0001
Skid_Resistance		1	0.2153	0.1048	4.2163	0.04

Injury Severity Regression Model
(for Driver Involvements in Angle Crashes at Unsignalized Intersections and Road Segments)

Parameter		DF	Estimate	Standard Error	Wald chi-square	Pr > chi-square
Intercept		1	-1.6527	0.2929	31.8351	<.0001
Driver_Ageg_Group_x	5	1	0.7545	0.1570	23.1013	<.0001
Driver_Ageg_Group_x	4	1	0.6348	0.1087	34.0811	<.0001
Driver_Ageg_Group_x	3	1	-0.0384	0.1157	0.1101	0.74
Driver_Ageg_Group_x	2	1	-0.0640	0.1230	0.2711	0.6026
Gender_x		1	0.3395	0.0771	19.4130	<.0001
Safety_Equipment_x	1	1	-1.1925	0.1036	132.5249	<.0001
Speeding_x	2	1	-0.4702	0.1455	10.4413	0.0012
Speeding_x	1	1	-0.5226	0.0895	34.1280	<.0001
Contributing_Cause_x	4	1	0.6367	0.1158	30.2482	<.0001
Contributing_Cause_x	3	1	1.3620	0.3547	14.7456	0.0001
Contributing_Cause_x	2	1	0.7418	0.1152	41.4444	<.0001
At_Fault_driver_x		1	-0.9381	0.1006	86.9271	<.0001
Ejected_x		1	1.4175	0.2020	49.2481	<.0001
Off_Roadway		1	0.5724	0.2543	5.0651	0.0244
point_impact_x		1	0.6128	0.0840	53.2458	<.0001
Type_of_Vehicle_x	5	1	-0.2032	0.3808	0.2847	0.5936
Type_of_Vehicle_x	4	1	0.0425	0.2237	0.0360	0.8495
Type_of_Vehicle_x	3	1	-1.3469	0.3481	14.9759	0.0001
Type_of_Vehicle_x	2	1	-0.2883	0.0947	9.2703	0.0023
nrural_urban		1	-0.4542	0.0796	32.5285	<.0001
Intersection		1	0.3333	0.0780	18.2468	<.0001
Speed_limit_x	1	1	-0.3904	0.0831	22.0720	<.0001
Access_class	9	1	-0.1155	0.1050	1.2094	0.2714
Access_class	7	1	-0.4794	0.1774	7.3042	0.0069
Access_class	6	1	-0.1724	0.1366	1.5926	0.207
Access_class	5	1	-0.2852	0.0953	8.9537	0.0028
Vision_Obstructed	4	1	-0.1805	0.2797	0.4164	0.5187
Vision_Obstructed	3	1	0.2046	0.1533	1.7802	0.1821
Vision_Obstructed	2	1	0.5592	0.1890	8.7565	0.0031

Injury Severity Regression Model
(for Driver Involvements in Angle Crashes on Road Segments)

Parameter		DF	Estimate	Standard Error	Wald chi-square	Pr > chi-square
Intercept		1	-1.1764	0.2316	25.8075	<.0001
Driver_Ageg_Group_x	**5**	1	0.4013	0.3171	1.6011	0.2057
Driver_Ageg_Group_x	**4**	1	0.6345	0.1846	11.8165	0.0006
Driver_Ageg_Group_x	**3**	1	-0.1709	0.1814	0.8879	0.3461
Driver_Ageg_Group_x	**2**	1	-0.2035	0.2072	0.9647	0.326
Gender_x		1	0.2763	0.1290	4.5863	0.0322
Safety_Equipment_x	**1**	1	-1.2254	0.1701	51.8960	<.0001
Speeding_x	**2**	1	-0.1402	0.2108	0.4421	0.5061
Speeding_x	**1**	1	-0.6231	0.1396	19.9152	<.0001
Contributing_Cause_x	**4**	1	0.6149	0.1734	12.5759	0.0004
Contributing_Cause_x	**3**	1	0.8473	0.5462	2.4065	0.1208
Contributing_Cause_x	**2**	1	0.4525	0.1882	5.7823	0.0162
At_Fault_driver_x		1	-0.6529	0.1643	15.7910	<.0001
Ejected_x		1	1.2973	0.2903	19.9694	<.0001
point_impact_x		1	0.3934	0.1436	7.5065	0.0061
Type_of_Vehicle_x	**5**	1	-0.3107	0.6707	0.2146	0.6432
Type_of_Vehicle_x	**4**	1	0.0415	0.3233	0.0164	0.898
Type_of_Vehicle_x	**3**	1	-1.4686	0.5251	7.8220	0.0052
Type_of_Vehicle_x	**2**	1	-0.3208	0.1598	4.0311	0.0447
nrural_urban		1	-0.4065	0.1292	9.9020	0.0017
Speed_limit_x	**1**	1	-0.2984	0.1418	4.4265	0.0354
Vision_Obstructed	**4**	1	-0.6491	0.6432	1.0185	0.3129
Vision_Obstructed	**3**	1	0.3976	0.2497	2.5357	0.1113
Vision_Obstructed	**2**	1	1.2684	0.2453	26.7284	<.0001

Injury Severity Regression Model
(for Driver Involvements in All Left turn Crashes)

Parameter		DF	Estimate	Standard Error	Wald chi-square	Pr > chi-square
Intercept		1	-2.5220	0.5620	20.1381	<.0001
Driver_Ageg_Group_x	**5**	1	0.7413	0.1736	18.2340	<.0001
Driver_Ageg_Group_x	**4**	1	0.4626	0.1316	12.3521	0.0004
Driver_Ageg_Group_x	**3**	1	-0.1562	0.1487	1.1041	0.2934
Driver_Ageg_Group_x	**2**	1	-0.2783	0.1610	2.9897	0.0838
Gender_x		1	0.4330	0.0955	20.5599	<.0001
Safety_Equipment_x	**1**	1	-1.1882	0.1287	85.2174	<.0001
At_Fault_driver_x		1	-0.5476	0.0953	33.0194	<.0001
Ejected_x		1	1.7024	0.3120	29.7663	<.0001
Off_Roadway		1	1.3012	0.5207	6.2457	0.0124
point_impact_x		1	0.4245	0.1230	11.9062	0.0006
Type_of_Vehicle_x	**5**	1	-1.1163	0.7713	2.0947	0.1478
Type_of_Vehicle_x	**4**	1	-0.5592	0.4039	1.9172	0.1662
Type_of_Vehicle_x	**3**	1	-1.1532	0.4460	6.6845	0.0097
Type_of_Vehicle_x	**2**	1	-0.4315	0.1232	12.2709	0.0005
Median_Type	**2**	1	-0.3607	0.1576	5.2373	0.0221
Speed_limit_x	**1**	1	-0.4257	0.1059	16.1603	<.0001
nAVGTFACT		1	0.0283	0.0114	6.1675	0.013
Skid_Resistance		1	0.2048	0.1020	4.0348	0.0446
LIGHTCDE	**Y**	1	0.4156	0.1717	5.8621	0.0155
LIGHTCDE	**P**	1	-0.1645	0.1627	1.0225	0.3119
Access_class	**9**	1	0.2216	0.1350	2.6947	0.1007
Access_class	**7**	1	-0.8372	0.2586	10.4775	0.0012
Access_class	**6**	1	-0.0858	0.2122	0.1635	0.6859
Access_class	**5**	1	-0.1263	0.1177	1.1519	0.2832

Injury Severity Regression Model
(for Driver Involvements in Left turn Crashes at Signalized Intersections)

Parameter		DF	Estimate	Standard Error	Wald chi-square	Pr > chi-square
Intercept		1	-0.8287	0.2662	9.6884	0.0019
Driver_Ageg_Group_x	5	1	0.5869	0.2597	5.1059	0.0238
Driver_Ageg_Group_x	4	1	0.4357	0.1971	4.8876	0.027
Driver_Ageg_Group_x	3	1	-0.1036	0.2146	0.2330	0.6293
Driver_Ageg_Group_x	2	1	-0.2946	0.2419	1.4832	0.2233
Gender_x		1	0.4606	0.1412	10.6489	0.0011
Safety_Equipment_x	1	1	-1.1705	0.1944	36.2387	<.0001
At_Fault_driver_x		1	-0.6279	0.1394	20.2978	<.0001
Ejected_x		1	1.9090	0.4276	19.9284	<.0001
Type_of_Vehicle_x	5	1	-12.9334	519.3	0.0006	0.9801
Type_of_Vehicle_x	4	1	-0.7414	0.5932	1.5623	0.2113
Type_of_Vehicle_x	3	1	-2.0175	1.0146	3.9537	0.0468
Type_of_Vehicle_x	2	1	-0.4209	0.1823	5.3272	0.021
Speed_limit_x	1	1	-0.4102	0.1617	6.4303	0.0112
Access_class	9	1	0.3932	0.2033	3.7395	0.0531
Access_class	7	1	-1.5053	0.5251	8.2187	0.0041
Access_class	6	1	0.0267	0.2614	0.0104	0.9187
Access_class	5	1	-0.1405	0.1693	0.6889	0.4065

Injury Severity Regression Model
(for Driver Involvements in Left turn Crashes at Unsignalized Intersections)

Parameter		DF	Estimate	Standard Error	Wald chi-square	Pr > chi-square
Intercept		1	-1.9603	0.3309	35.0867	<.0001
Driver_Ageg_Group_x	5	1	0.9643	0.1846	27.2879	<.0001
Driver_Ageg_Group_x	4	1	0.6824	0.1354	25.4021	<.0001
Driver_Ageg_Group_x	3	1	0.0481	0.1506	0.1020	0.7494
Driver_Ageg_Group_x	2	1	0.0218	0.1543	0.0200	0.8876
Gender_x		1	0.4241	0.0956	19.6901	<.0001
Safety_Equipment_x	1	1	-1.1851	0.1290	84.3853	<.0001
Speeding_x	2	1	-0.7957	0.2043	15.1755	<.0001
Speeding_x	1	1	-0.4932	0.1182	17.4151	<.0001
Contributing_Cause_x	4	1	0.6149	0.1589	14.9704	0.0001
Contributing_Cause_x	3	1	1.7911	0.4703	14.5039	0.0001
Contributing_Cause_x	2	1	0.9028	0.1504	36.0398	<.0001
At_Fault_driver_x		1	-1.1768	0.1295	82.6209	<.0001
Ejected_x		1	1.6846	0.2142	61.8763	<.0001
point_impact_x		1	0.7471	0.1053	50.3600	<.0001
Private_vehicle_use_		1	0.6991	0.2747	6.4741	0.0109
nrural_urban		1	-0.5700	0.0993	32.9765	<.0001
Speed_limit_x	1	1	-0.4737	0.0995	22.6644	<.0001
Skid_Resistance		1	0.2444	0.1057	5.3454	0.0208
AUX_Lane_Num	3	1	0.1856	0.1799	1.0639	0.3023
AUX_Lane_Num	2	1	-0.3696	0.1617	5.2268	0.0222
AUX_Lane_Num	1	1	0.1951	0.1053	3.4316	0.064
Type_Friction_Course	9	1	0.1983	0.1174	2.8520	0.0913
Type_Friction_Course	5	1	0.4085	0.2168	3.5489	0.0596
Type_Friction_Course	4	1	0.0185	0.1193	0.0241	0.8766
Type_Friction_Course	1	1	-0.3609	0.2340	2.3786	0.123

Injury Severity Regression Model
(for Driver Involvements in Left turn Crashes at Unsignalized Intersections and Road Segments)

Parameter		DF	Estimate	Standard Error	Wald chi-square	Pr > chi-square
Intercept		1	-2.3565	0.6209	14.4062	0.0001
Driver_Ageg_Group_x	**5**	1	0.9031	0.2320	15.1473	<.0001
Driver_Ageg_Group_x	**4**	1	0.5620	0.1756	10.2410	0.0014
Driver_Ageg_Group_x	**3**	1	-0.1726	0.2089	0.6830	0.4085
Driver_Ageg_Group_x	**2**	1	-0.1906	0.2129	0.8015	0.3707
Gender_x		1	0.5075	0.1269	16.0008	<.0001
Safety_Equipment_x	**1**	1	-1.1760	0.1710	47.3054	<.0001
At_Fault_driver_x		1	-0.4866	0.1333	13.3262	0.0003
Ejected_x		1	1.3939	0.3379	17.0162	<.0001
Off_Roadway		1	1.3493	0.5954	5.1364	0.0234
point_impact_x		1	0.5528	0.1482	13.9205	0.0002
nrural_urban		1	-0.3368	0.1294	6.7732	0.0093
Median_Type	**2**	1	-0.6033	0.1703	12.5484	0.0004
Speed_limit_x	**1**	1	-0.5336	0.1329	16.1343	<.0001

Injury Severity Regression Model
(for Driver Involvements in Left turn Crashes at Road Segments)

Parameter		DF	Estimate	Standard Error	Wald chi-square	Pr > chi-square
Intercept		1	-2.8431	0.4916	33.4455	<.0001
Gender_x		1	0.9332	0.2414	14.9487	0.0001
Safety_Equipment_x	1	1	-0.9833	0.3034	10.5069	0.0012
Ejected_x		1	2.1826	0.6625	10.8553	0.001
point_impact_x		1	0.8141	0.2692	9.1460	0.0025
Access_class	9	1	-0.0520	0.3133	0.0275	0.8683
Access_class	7	1	-1.7441	0.7461	5.4647	0.0194
Access_class	6	1	-0.9079	0.3747	5.8707	0.0154
Access_class	5	1	-0.6067	0.2942	4.2533	0.0392
Day_of_Week		1	0.5676	0.2635	4.6414	0.0312

Injury Severity Regression Model
(for Driver Involvements in All Fixed Object Crashes)

Parameter		DF	Estimate	Standard Error	Wald chi-square	Pr > chi-square
Intercept		1	-0.7738	0.1778	18.9478	<.0001
Safety_Equipment_x	1	1	-0.9289	0.0970	91.6141	<.0001
Speeding_x	2	1	0.1689	0.1422	1.4115	0.2348
Speeding_x	1	1	-0.5416	0.1074	25.4412	<.0001
Contributing_Cause_x	4	1	0.8446	0.1264	44.6827	<.0001
Contributing_Cause_x	3	1	0.5698	0.2340	5.9290	0.0149
Contributing_Cause_x	2	1	0.6256	0.2158	8.4064	0.0037
At_Fault_driver_x		1	-0.8076	0.0996	65.7982	<.0001
Ejected_x		1	1.3724	0.1696	65.4804	<.0001
Vehicle_Maneuver_x	4	1	0.2137	0.1512	1.9976	0.1576
Vehicle_Maneuver_x	3	1	-0.8341	0.1992	17.5245	<.0001
Vehicle_Maneuver_x	2	1	0.4623	0.3331	1.9261	0.1652
ADT_PER_LANE		1	-0.0904	0.0152	35.1822	<.0001

Injury Severity Regression Model
(for Driver Involvements in Fixed Object Crashes at Signalized Intersections)

Parameter		DF	Estimate	Standard Error	Wald chi-square	Pr > chi-square
Intercept		1	-2.2217	0.2150	106.7891	<.0001
Speeding_x	2	1	0.9660	0.3557	7.3744	0.0066
Speeding_x	1	1	-0.1802	0.2918	0.3815	0.5368
point_impact_x		1	1.0473	0.3135	11.1600	0.0008

Injury Severity Regression Model
(for Driver Involvements in Fixed Object Crashes at Unsignalized Intersections)

Parameter		DF	Estimate	Standard Error	Wald chi-square	Pr > chi-square
Intercept		1	-1.3481	0.3765	12.8212	0.0003
driver_ageg_group_x	4	1	1.2549	0.4199	8.9321	0.0028
driver_ageg_group_x	3	1	0.0612	0.3575	0.0293	0.8641
driver_ageg_group_x	2	1	-0.2267	0.3831	0.3501	0.5541
gender_x		1	-0.7600	0.2781	7.4669	0.0063
Safety_Equipment_x	1	1	-0.9700	0.3077	9.9399	0.0016
Speeding_x	2	1	-0.2839	0.4756	0.3562	0.5507
Speeding_x	1	1	-1.0386	0.3256	10.1768	0.0014
Ejected_x		1	1.7415	0.4909	12.5855	0.0004
Off_Roadway		1	0.5897	0.2694	4.7926	0.0286
roadway_curve		1	0.9650	0.4188	5.3101	0.0212

Injury Severity Regression Model
(for Driver Involvements in Fixed Object Crashes at Unsignalized Intersections and Road Segments)

Parameter		DF	Estimate	Standard Error	Wald chi-square	Pr > chi-square
Intercept		1	-0.5106	0.2271	5.0567	0.0245
driver_ageg_group_x	4	1	0.4100	0.1894	4.6882	0.0304
driver_ageg_group_x	3	1	-0.1661	0.1250	1.7645	0.1841
driver_ageg_group_x	2	1	-0.3107	0.1413	4.8382	0.0278
gender_x		1	-0.2148	0.1061	4.0985	0.0429
Safety_Equipment_x	1	1	-0.9985	0.1069	87.2398	<.0001
Speeding_x	2	1	0.0596	0.1593	0.1400	0.7083
Speeding_x	1	1	-0.6351	0.1216	27.3001	<.0001
Contributing_Cause_x	4	1	0.9551	0.1449	43.4477	<.0001
Contributing_Cause_x	3	1	0.8290	0.2534	10.7035	0.0011
Contributing_Cause_x	2	1	0.7178	0.2355	9.2853	0.0023
At_Fault_driver_x		1	-0.7611	0.1081	49.5541	<.0001
Ejected_x		1	2.5266	0.4413	32.7783	<.0001
gender_x*Ejected_x		1	-1.3333	0.4783	7.7692	0.0053
Vehicle_Maneuver_x	4	1	0.2261	0.1623	1.9417	0.1635
Vehicle_Maneuver_x	3	1	-1.0755	0.2713	15.7155	<.0001
Vehicle_Maneuver_x	2	1	0.2499	0.4788	0.2724	0.6017
nrural_urban		1	-0.3416	0.1018	11.2468	0.0008
roadway_curve		1	0.4823	0.1390	12.0386	0.0005
nType_of_Shoulder	3	1	0.0894	0.1195	0.5605	0.4541
nType_of_Shoulder	2	1	-0.2208	0.1351	2.6704	0.1022
ADT_PER_LANE		1	-0.0958	0.0172	31.1202	<.0001
LIGHTCDE	Y	1	0.4991	0.1796	7.7260	0.0054
LIGHTCDE	P	1	0.0464	0.1771	0.0687	0.7933

Injury Severity Regression Model
(for Driver Involvements in Fixed Object Crashes at Road Segments)

Parameter		DF	Estimate	Standard Error	Wald chi-square	Pr > chi-square
Intercept		1	-1.1679	0.4021	8.4356	0.0037
driver_ageg_group_x	4	1	0.1358	0.2136	0.4041	0.525
driver_ageg_group_x	3	1	-0.2216	0.1336	2.7516	0.0972
driver_ageg_group_x	2	1	-0.3585	0.1520	5.5616	0.0184
Safety_Equipment_x	1	1	-0.9725	0.1134	73.5886	<.0001
Speeding_x	2	1	0.0983	0.1695	0.3368	0.5617
Speeding_x	1	1	-0.5293	0.1303	16.5059	<.0001
Contributing_Cause_x	4	1	0.7839	0.1889	17.2224	<.0001
Contributing_Cause_x	3	1	0.6254	0.2932	4.5506	0.0329
Contributing_Cause_x	2	1	0.6612	0.2681	6.0820	0.0137
At_Fault_driver_x		1	-0.8135	0.1149	50.0888	<.0001
Ejected_x		1	1.4156	0.1992	50.5008	<.0001
Vehicle_Maneuver_x	4	1	0.3220	0.1666	3.7382	0.0532
Vehicle_Maneuver_x	3	1	-0.8227	0.3393	5.8803	0.0153
Vehicle_Maneuver_x	2	1	0.3856	0.5500	0.4916	0.4832
Private_vehicle_use_		1	0.8087	0.3275	6.0969	0.0135
nrural_urban		1	-0.3011	0.1091	7.6188	0.0058
Multivehicle	1	1	-0.4074	0.1760	5.3587	0.0206
roadway_curve		1	0.3927	0.1470	7.1369	0.0076
nType_of_Shoulder	3	1	0.1265	0.1279	0.9777	0.3228
nType_of_Shoulder	2	1	-0.1971	0.1427	1.9071	0.1673
ADT_PER_LANE		1	-0.1005	0.0183	30.0577	<.0001

APPENDIX F: INJURY SEVERITY MODELS' FINAL ANALYSIS – MODELS BY LAND USE

Injury Severity Regression Model
(for Driver Involvements Rural Area Crashes)

Parameter		DF	Estimate	Standard Error	Wald chi-square	Pr > chi-square
Intercept		1	-0.5043	0.1603	9.8998	0.0017
Driver_Ageg_Group_x	5	1	0.6002	0.0996	36.3505	<.0001
Driver_Ageg_Group_x	4	1	0.3451	0.0634	29.6521	<.0001
Driver_Ageg_Group_x	3	1	-0.2329	0.0565	17.0009	<.0001
Driver_Ageg_Group_x	2	1	-0.2731	0.0625	19.1198	<.0001
Ejected_x		1	1.5241	0.0971	246.5351	<.0001
Speeding_x	2	1	-0.2914	0.1683	2.9983	0.0834
Speeding_x	1	1	-0.7995	0.1132	49.9050	<.0001
Gender_x		1	0.2283	0.0862	7.0192	0.0081
Safety_Equipment_x	1	1	-1.3298	0.0617	464.2625	<.0001
Gender_x*Safety_Equi	1	1	0.2284	0.0957	5.6989	0.017
At_Fault_driver_x		1	-0.6136	0.0519	139.7425	<.0001
Physical_Defects_x		1	0.3742	0.0991	14.2670	0.0002
Harmful_Event_Group_	7	1	0.1112	0.0598	3.4585	0.0629
Harmful_Event_Group_	6	1	0.4262	0.0939	20.6108	<.0001
Harmful_Event_Group_	5	1	-0.3808	0.1298	8.6110	0.0033
Harmful_Event_Group_	4	1	0.7102	0.0819	75.2792	<.0001
Harmful_Event_Group_	3	1	0.5817	0.0630	85.1846	<.0001
Harmful_Event_Group_	2	1	1.1362	0.1184	92.0334	<.0001
Contributing_Cause_x	4	1	0.4955	0.0545	82.5279	<.0001
Contributing_Cause_x	3	1	0.5854	0.1474	15.7788	<.0001
Contributing_Cause_x	2	1	0.4755	0.0677	49.2807	<.0001
Type_of_Vehicle_x	5	1	-0.2969	0.1881	2.4910	0.1145
Type_of_Vehicle_x	4	1	-0.2579	0.1097	5.5226	0.0188
Type_of_Vehicle_x	3	1	-1.1844	0.1463	65.5539	<.0001
Type_of_Vehicle_x	2	1	-0.1949	0.0458	18.0773	<.0001
point_impact_x		1	0.0347	0.0958	0.1308	0.7176
point_imp*Speeding_x	2	1	0.2205	0.1932	1.3020	0.2538
point_imp*Speeding_x	1	1	0.3656	0.1186	9.5046	0.002
Off_Roadway		1	-0.5274	0.1020	26.7339	<.0001
Off_Roadw*Speeding_x	2	1	-0.1372	0.1767	0.6031	0.4374
Off_Roadw*Speeding_x	1	1	0.2004	0.1210	2.7426	0.0977
Off_Roadw*Multivehic	1	1	0.7159	0.1543	21.5374	<.0001
nWork_Area_x	3	1	-0.2917	0.1364	4.5752	0.0324

Parameter		DF	Estimate	Standard Error	Wald chi-square	Pr > chi-square
nWork_Area_x	**2**	1	-0.2998	0.1208	6.1603	0.0131
Multivehicle	**1**	1	-0.7436	0.1377	29.1570	<.0001
Intersect*Multivehic	**1**	1	0.1941	0.0493	15.4788	<.0001
Speed_limit_x	**1**	1	-0.4105	0.0456	80.9731	<.0001
ADT_PER_LANE		1	-0.0290	0.00773	14.0392	0.0002
nAVGTFACT		1	0.00950	0.00376	6.3914	0.0115
Traffic_Control	**3**	1	0.1195	0.0649	3.3901	0.0656
Traffic_Control	**2**	1	-0.0325	0.0508	0.4098	0.522
Access_class1	**9**	1	0.0869	0.0496	3.0770	0.0794
Access_class1	**7**	1	0.1101	0.1091	1.0187	0.3128
Access_class1	**6**	1	-0.0360	0.0883	0.1666	0.6832
Access_class1	**5**	1	-0.0904	0.0518	3.0405	0.0812
nType_of_Shoulder	**3**	1	-0.0203	0.0535	0.1443	0.7041
nType_of_Shoulder	**2**	1	-0.1203	0.0475	6.4130	0.0113
Lane_width	**4**	1	-0.1214	0.0691	3.0892	0.0788
Lane_width	**3**	1	-0.3965	0.1497	7.0152	0.0081
Lane_width	**2**	1	-0.2186	0.0800	7.4690	0.0063
roadway_curve		1	0.1786	0.0798	5.0155	0.0251
Sidewalk_width_group	**3**	1	-0.3018	0.0626	23.2094	<.0001
Sidewalk_width_group	**2**	1	-0.2264	0.0513	19.4807	<.0001
LIGHTCDE	**Y**	1	-0.1070	0.0872	1.5061	0.2197
LIGHTCDE	**P**	1	-0.3193	0.0724	19.4568	<.0001
Skid_Resistance		1	0.1636	0.0420	15.1452	<.0001
Day_of_Week		1	-0.1028	0.0427	5.8031	0.016

Injury Severity Regression Model
(for Driver Involvements Urban Area Crashes)

Parameter		DF	Estimate	Standard Error	Wald chi-square	Pr > chi-square
Intercept		1	-1.5993	0.2344	46.5348	<.0001
Driver_Ageg_Group_x	5	1	0.3323	0.1231	7.2812	0.007
Driver_Ageg_Group_x	4	1	0.3623	0.0737	24.1588	<.0001
Driver_Ageg_Group_x	3	1	-0.2811	0.0645	19.0094	<.0001
Driver_Ageg_Group_x	2	1	-0.2467	0.0720	11.7458	0.0006
Ejected_x		1	1.3682	0.1010	183.4125	<.0001
Speeding_x	2	1	0.0365	0.1749	0.0435	0.8348
Speeding_x	1	1	-1.0925	0.1591	47.1420	<.0001
Gender_x		1	0.1738	0.0946	3.3762	0.0661
Safety_Equipment_x	1	1	-1.0277	0.0714	207.0976	<.0001
Gender_x*Safety_Equi	1	1	0.1953	0.1061	3.3886	0.0656
At_Fault_driver_x		1	-0.6477	0.0586	122.1641	<.0001
Residence_Code_x		1	0.2681	0.1144	5.4929	0.0191
Physical_Defects_x		1	0.4966	0.1127	19.4027	<.0001
Harmful_Event_Group_	7	1	0.0623	0.0646	0.9286	0.3352
Harmful_Event_Group_	6	1	0.8647	0.1070	65.3435	<.0001
Harmful_Event_Group_	5	1	-0.1242	0.1254	0.9806	0.3221
Harmful_Event_Group_	4	1	0.9320	0.0876	113.3232	<.0001
Harmful_Event_Group_	3	1	0.5885	0.0704	69.8856	<.0001
Harmful_Event_Group_	2	1	1.0135	0.1209	70.2737	<.0001
Contributing_Cause_x	4	1	0.4344	0.0590	54.1346	<.0001
Contributing_Cause_x	3	1	0.5039	0.1706	8.7235	0.0031
Contributing_Cause_x	2	1	0.4626	0.0733	39.7874	<.0001
Type_of_Vehicle_x	5	1	-0.2312	0.2023	1.3061	0.2531
Type_of_Vehicle_x	4	1	0.4216	0.1144	13.5778	0.0002
Type_of_Vehicle_x	3	1	-0.7868	0.1774	19.6681	<.0001
Type_of_Vehicle_x	2	1	-0.1978	0.0546	13.1418	0.0003
point_impact_x		1	0.1569	0.1118	1.9695	0.1605
point_imp*Speeding_x	2	1	0.1902	0.1913	0.9887	0.32
point_imp*Speeding_x	1	1	0.3291	0.1353	5.9144	0.015
Off_Roadway		1	-0.4003	0.1308	9.3661	0.0022
Off_Roadw*Speeding_x	2	1	-0.3837	0.1848	4.3133	0.0378
Off_Roadw*Speeding_x	1	1	0.4144	0.1663	6.2050	0.0127
Off_Roadw*Multivehic	1	1	0.6326	0.1897	11.1170	0.0009

Parameter		DF	Estimate	Standard Error	Wald chi-square	Pr > chi-square
nWork_Area_x	**3**	1	-0.00298	0.1590	0.0004	0.985
nWork_Area_x	**2**	1	-0.2582	0.1311	3.8816	0.0488
Multivehicle	**1**	1	-0.6568	0.1763	13.8756	0.0002
Intersect*Multivehic	**1**	1	0.5522	0.1456	14.3812	0.0001
Speed_limit_x	**1**	1	-0.3547	0.0570	38.7016	<.0001
ADT_PER_LANE		1	-0.0243	0.00877	7.7165	0.0055
nAVGTFACT		1	0.0133	0.00619	4.5806	0.0323
LIGHTING	**Y**	1	0.2939	0.1335	4.8449	0.0277
LIGHTING	**P**	1	1.1878	0.4612	6.6322	0.01
Access_class1	**9**	1	-0.00889	0.0779	0.0130	0.9092
Access_class1	**7**	1	-0.4196	0.0944	19.7714	<.0001
Access_class1	**6**	1	-0.2424	0.0797	9.2445	0.0024
Access_class1	**5**	1	-0.1416	0.0556	6.4838	0.0109
nType_of_Shoulder	**3**	1	0.2251	0.0539	17.4334	<.0001
nType_of_Shoulder	**2**	1	0.1165	0.0680	2.9373	0.0866
Lane_width	**4**	1	-0.2144	0.0710	9.1189	0.0025
Lane_width	**3**	1	-0.1445	0.0816	3.1337	0.0767
Lane_width	**2**	1	-0.1581	0.0749	4.4504	0.0349
roadway_curve		1	0.4289	0.1083	15.6718	<.0001
LIGHTCDE	**Y**	1	0.2658	0.0703	14.2873	0.0002
LIGHTCDE	**P**	1	-0.0534	0.0713	0.5596	0.4544
Type_Friction_Course	**9**	1	-0.2195	0.0556	15.5798	<.0001
Type_Friction_Course	**5**	1	-0.2163	0.1342	2.5984	0.107
Type_Friction_Course	**4**	1	-0.2435	0.0617	15.5718	<.0001
Type_Friction_Course	**1**	1	-0.4768	0.0950	25.2035	<.0001
Intersection		1	-0.4626	0.1375	11.3234	0.0008
Skid_Resistance		1	0.1188	0.0468	6.4412	0.0112
Day_of_Week		1	-0.0959	0.0483	3.9334	0.0473

LIST OF REFERENCES

Abdel-Aty, M. "Analysis of driver injury severity levels at multiple locations using ordered probit models." Journal of Safety Research, Vol. 34, No. 5, 597-603, 2003.

Abdel-Aty, M. and H. Abdelwahab. "Configuration analysis of two-vehicle rear-end crashes." Transportation Research Record, No. 1840, 140-147, 2003.

Abdel-Aty, M. and H. Abdelwahab. "Analysis and prediction of traffic fatalities resulting from angle collisons including the effect of vehicles' configuration and compatibility." Accident Analysis & Prevention, Volume 36, 457-469, 2003.

Abdel-Aty , M., C. Chen and E. Radwan. "Analsyis of crash-involvement trends by drivers' age in Florida." ITE Journal on the Web, February 1999, 69-73.

Abdel-Aty, M., C. Chen and J. Schott. "An assessment of the effect of driver age on traffic accident involvement using log-linear models" Accident Analysis & Prevention, Volume 30, No. 6, 851-861, 1998.

Abdel-Aty, M. and E. Radwan. "Modeling traffic accident occurrence and involvement." Accident Analysis & Prevention, Volume 32, No. 5, 633-642, 2000.

Abdel-Aty, M., H. Salkapuram, C. Lee, and P. Brady. "A Simplistic, Practical approach to identify Traffic Crash Profiles at Signalized Intersections." ITE Journal on the Web, April 2006, 28-33.

Abdel-Aty, M. and J. Keller. "Exploring the overall and specific crash severity levels at signalized intersections." Accident Analysis & Prevention, Volume 37, No. 3, 417-425, 2005.

Abdel-Aty, M. and X. Wang. "Crash Estimation at Signalized Intersections Along Corridors: Analyzing Spatial Effect and Identifying Significant Factors." Transportation Research Record, No. 1953, 98-111, 2006.

Al-Ghamdi, A. S., “Using logistic regression to estimate the influence of accident factors on accident severity.” Accident Analysis & Prevention, Vol. 34, 729-741, 2002.

American Association of State Highway and Transportation Officials. “Transportation Invest in our Future: Surface Transportation Policy Recommendations for the National Surface Transportation Policy and Revenue Study Commission.” AASHTO, Washington DC, 2007.

Bauer, K. M. and D. W. Harwood. “Statistical Models for At-grade Intersection Accidents.” Publication FHWA-RD-96-125. FHWA, U.S. Department of Transportation, 1996.

Bédard, M., G. H. Guyatt, M. J. Stones and J. P. Hirdes. “The independent contribution of driver, crash, and vehicle characteristics to driver fatalities.” Accident Analysis & Prevention, Volume 34, No. 6, 717-727, 2002.

Berhanu, G. “Models relating traffic safety with road environment and traffic flows on arterial roads in Addis Ababa.” Accident Analysis & Prevention, Volume 36, No. 5, 697-704, 2004.

Brown, H.C., A.P. Tarko . “Effects of Access Control on Safety on Urban Arterial Streets.” Transportation Research Record, No. 1665, 68-74, 1999.

Bonneson, J. A. and P. T. McCoy. “Effect of Median Treatment on Urban Arterial Safety: An Accident Prediction Model.” Transportation Research Record, No. 1581, 27-36, 1997.

Chang, L. and F. Mannering. “Analysis of injury severity and vehicle occupancy in truck- and non-truck-involved accidents.” Accident Analysis & Prevention , Volume 31, No. 5, 579-592, 1999.

Dahir, S. and L.G. Wade. “Wet-Pavement Safety Programs. NCHRP Synthesis of Highway Practice 158.” Transportation Research Board, National Research Council, Washington, DC, 1990.

Devon, E. L. and C. G. Hampton. “The fatality and injury risk of light truck impacts with pedestrians in the United States.” Accident Analysis & Prevention, Volume 36, No. 2, 295-304, 2004.

Dissanayake, S. and J. J. Lu. “Factors influential in making an injury severity difference to older drivers involved in fixed object–passenger car crashes.” Accident Analysis & Prevention, Vol. 34, No. 5, 609-618, 2002.

Duncan, C. S., A. J. Khattak, and F. M . Council. “Applying the Ordered Probit Model to Injury Severity in Truck-Passenger Car Rear-End Collisions.” Transportation Research Record, No. 1635, 8-11, 1998.

Toy, E. L. and J. K. Hammitt. “Safety Impacts of SUVs, Vans, and Pickup Trucks in Two-Vehicle Crashes.” Risk Analysis, Volume 23, No. 4, 641-650, 2003.

Eisele, W.L. and W. E. Frawley. "Estimating the Safety and Operational Impact of Raised Medians and Driveway Density: Experiences from Texas and Oklahoma Case Studies." Transportation Research Record, No. 1931, 108-116, 2005.

Elvik, R. and A. B. Mysen. “Incomplete Accident Reporting: Meta-Analysis of Studies Made in 13 Countries.” Transportation Research Record, No. 1665, 133-140, 1999.

Elvik, R. and P. Greibe. "Road safety effects of porous asphalt: a systematic review of evaluation studies." Accident Analysis & Prevention, Vol. 37, 515-522, 2005.

Florida Department of Highway Safety and Motor Vehicles. "Florida Traffic Crash Records Database: 2002 to Present." Office of Management Research and Development. Tallahassee, FL, 2004.

Florida Department of Highway Safety and Motor Vehicles. "Traffic Crash Statistics 2005." Office of Management Research and Development. Tallahassee, FL, 2007.

Florida Department of Highway Safety and Motor Vehicles. "Traffic Crash Statistics Report 2006: A Compilation of Motor Vehicle Crash Data from the Florida Crash Records Database." http://www.flhsmv.gov/hsmvdocs/CS2006.pdf Accessed January 9, 2008.

Florida Department of Transportation. "Access Management: Balancing Access and Mobility." Office of the State Transportation Planner, Systems Planning Office, http://www.dot.state.fl.us/planning/systems/sm/accman/pdfs/ampromo3.pdf Accessed November 13, 2007.

Florida Department of Transportation. "Florida Reaches Highest Safety Belt Usage Rate in State's History." Public Information Office. http://www.dot.state.fl.us/publicinformationoffice/moreDOT/spenews/2004/safetybelt.htm Accessed February 19, 2008.

Florida Department of Transportation. "Highway Mileage Reports." Office of the State Transportation Planner, Transportation Statistics Office, http://www.dot.state.fl.us/planning/statistics/mileage-rpts/ Accessed February 20, 2008.

Florida Department of Transportation. “Florida Strategic Highway Safety Plan. 2006.” http://www.dot.state.fl.us/Safety/posted%20documents/Florida%20Comp%20SHSP%20for%20print.pdf Accessed November 13, 2007.

Florida Department of Transportation. “Manual of Uniform Minimum Standards for Design, Construction and Maintenance for Streets and Highways (Commonly known as the "Florida Greenbook").” May 2005 Edition.

Florida Department of Transportation. “Transportation Statistics RCI Field Handbook.” Office of the State Transportation Planner, Transportation Statistics Office, http://www.dot.state.fl.us/planning/statistics/rci/fieldhandbook/fullrpt.pdf Accessed July 20, 2007.

Florida Department of Transportation. “User Manual for the Crash Analysis Reporting System (CAR).” Safety Office Crash Records Division. Tallahassee, FL. September 2005.

Gattis, J. L., R. Balakumar and L. K. Duncan. "Effects of Rural Highway Median Treatments and Access." Transportation Research Record, No. 1931, 99-107, 2005.

Hauer, E., J. Ng and J. Lovel. “Estimation of Safety at Signalized Intersections.” Transportation Research Record, No. 1185, 48-61, 1988.

Hauer, E., F. Council and Y. Mohammedshah. “Safety Models for Urban Four-Lane Undivided Road Segments.” Transportation Research Record, No. 1897, 96-105, 2006.

Hauer, E. and A. S. Hakkert. “Extent and Some Implications of Incomplete Accident Reporting.” Transportation Research Record, No. 1185, 1-10, 1988.

Hosmer, D. W., and S. Lemeshow. “Applied Logistic Regression.” John Wiley and Sons, New York, 2000.

Hunter, W. and J. Feaganes. “Effect of wide curb lanes conversions on bicycle and motor vehicle interactions” Report BA 784, Florida Department of Transportation. Tallahasse, FL, April 2006.

Islam, S. and F. Mannering. “Driver aging and its effect on male and female single-vehicle accident injuries: Some additional evidence.” Journal of Safety Research, Volume 37, No. 3, 267-276, 2006.

Jaccard, J. and T. Dodge. “Analyzing Contingent Effects in Regression Models.” In Handbook of Data Analysis, Edited by Hardy, M. and Bryman, A. Sage Publications, 2004.

James, J. and K. Kim. “Restraint use by children involved in crashes in Hawaii, 1986–1991.” Transportation Research Record, No. 1560, 8-11, 1996.

Karlaftis, M. G. and I. Golias. “Effects of road geometry and traffic volumes on rural roadway accident rates.” Accident Analysis & Prevention, Volume 34, No. 3, 357-365, 2002.

Kendall, M. G. and A. Stuart. “The advanced theory of statistics- Volume 2: Inference and Relationship.” England: Charles Griffin & Company Limited, 1979.

Kim, K., L. Nitz, J. Richardson, and L. Li. "Personal and behavioral predictors of automobile crash and injury severity." Accident Analysis & Prevention Volume 27, No. 4, 469-481, 1995.

Khattak, S. J. “Injury Severity in Multivehicle Rear-End Crashes.” Transportation Research Record, No. 1746, 59-68, 2001.

Khorashadi, A., D. Niemeier, V. Shankar and F. Mannering. “Differences in rural and urban driver-injury severities in accidents involving large-trucks: An exploratory analysis.” Accident Analysis & Prevention, Volume 37, No. 5, 910-921, 2005.

Kockelman, K. M. and J. Kweon "Driver Injury Severity: an application of ordered probit models." Accident Analysis & Prevention, Vol. 34, No. 1, 313-321, 2002.

Krull, K.A., A.J. Khattak and F. M. Council. "Injury Effects of Rollovers and Events Sequence in Single-Vehicle Crashes." Transportation Research Record, No. 1717, 46-54, 2000.

Le, C.T. "Applied Categorical Data Analysis." John Wiley and Sons, New York, 1998.

Lee, J. and F. Mannering. "Impact of roadside features on the frequency and severity of run-off-roadway accidents: an empirical analysis." Accident Analysis & Prevention, Volume 34, No. 2, 149-161, 2002.

Li, L., K. Kim and L. Nitz. "Predictors of safety belt use among crash-involved drivers and front seat passengers: adjusting for over-reporting." Accident Analysis & Prevention, Volume 31, No. 6, 631-638, 1999.

Liu, P., J. Lu and H. Chen. "Safety effects of the separation distances between driveway exits and downstream U-turn locations." Accident Analysis & Prevention, Volume 40, No. 2, 760-767, 2008.

McCoy, P. T. and M. S. Malone. "Safety Effects of Left-Turn Lanes on Urban Four-Lane Roadways." Transportation Research Record, No. 1239, 17-22, 1989.

Menard, S. "Applied Logistic Regression Analysis." Sage Publications, London, 2001.

Mercier, C. R., M. C. Shelley, J.B. Rimkus, and J. C. Mercier. "Age and Gender as Predictors of Injury Severity in Head-On Highway Vehicular Collisions." Transportation Research Record, No. 1581, 37-46, 1997.

Obeng, K. "Some determinants of possible injuries in crashes at signalized intersections." Journal of Safety Research, Volume 38, 103-112, 2007.

O'Donnell, C. J. and D. H. Connor. “Predicting the severity of motor vehicle accident injuries using models of ordered multiple choice.” Accident Analysis & Prevention, Volume 28, No. 6, 739-753, 1996.

Potts, I. B, D. W. Harwood, J. F. Ringert, and K. M. Bauer. “Operational and Safety Effects of Right-Turn Deceleration Lanes on Urban and Suburban Arterials." Presented at the Transportation Research Board 86th Annual Meeting, Washington, DC, 2007.

Potts, I. B, D. W. Harwood and K. R. Richard. “Relationship of Lane Width to Safety for Urban and Suburban Arterials.” Transportation Research Record, No. 2023, 63-82, 2007.

Roess, R. P., E. S. Prassas and W. R. McShane. “Traffic Engineering: Third Edition.” Pearson Prentice Hall, New Jersey, 2004.

Saccomanno, F. F., S. A, Nassar, and J. H. Shortreed. “Reliability of Statistical Road Accident Injury Severity Models.” Transportation Research Record, No. 1542, 14-23, 1996.

Shankar, V., F. Mannering and W. Barfield. “Statistical analysis of accident severity on rural freeways.”Accident Analysis & Prevention, Volume 28, No. 3, 391-401, 1996.

Shefer D. and P. Rietveld. “Congestion and Safety on Highways: Towards an Analytical Model.” Urban Studies, Vol. 34, No. 4, 679-692, 1997.

Simonoff, J. S. “Analyzing Categorical Data.” Springer, New York, 2003.

State Safety Office. “Skid Hazard Reporting System.” Florida Department of Transportation, Tallahassee, FL, 1995 (updated May 2006).

Strathman, J., K. Duecker, J. Zhang, and T. Williams. “Analysis of Design Attributes and Crashes on the Oregon Highway System.” Center for Urban Studies, Portland State University, for Oregon Department of Transportation Research Group and FHWA, U.S. Department of Transportation, 2001.

Streff, F. M. and A. C. Wagenaar “Are there really shortcuts?: Estimating seat belt use with self-report measures.” Accident Analysis & Prevention. Volume 21, No. 6, 509-516, 1989.

Sze, N.N. and S.C. Wong. “Diagnostic analysis of the logistic model for pedestrian injury severity in traffic crashes.” Accident Analysis & Prevention , Volume 39, No. 6, 1267-1278, 2007.

Tabachnick, B. G. and L.S. Fidell. “Using Multivariate Statistics: Fifth Edition.” Pearson/Allyn&Bacon. Boston, 2007.

Toy, E.L., J K. Hammitt. "Safety Impacts of SUVs, Vans, and Pickup Trucks in Two-Vehicle Crashes." Risk Analysis Volume 23, No. 4, 641-650, 2003.

Ulfarsson, G. F. and F. L. Mannering. “Differences in male and female injury severities in sport-utility vehicle, minivan, pickup and passenger car accidents.” Accident Analysis & Prevention, Volume 36, No. 2, 135-147, 2004.

Valent, F., F. Schiava, C. Savonitto, T. Gallo, S. Brusaferro and F. Barbone . “Risk factors for fatal road traffic accidents in Udine, Italy.” Accident Analysis & Prevention, Volume 34, 71-84, 2002.

Williams, K. M. “Cooperative Agreements for Corridor Management: A Synthesis of Highway Practice: NCHRP Synthesis 337.” Transportation Research Board of the National Academies, Washington, DC, 2004.

Wang, X., and M. Abdel-Aty. “Temporal and spatial analyses of rear-end crashes at signalized intersections.” Accident Analysis & Prevention, Volume 38, 1137-1150, 2006.

Wang, X., and M. Abdel-Aty. “Left-Turn crash injury severity analyses using partial proportional odds models: significant factors and varying effects.” Presented at the Transportation Research Board 87th Annual Meeting, Washington, DC, 2008.

Wang, X., M. Abdel-Aty, and P. A. Brady. “Crash Estimation at Signalized Intersections: Significant Factors and Temporal Effect.” Transportation Research Record: Journal of the Transportation Research Board, No. 1953, 10-20, 2006.

Wang, X., M. Abdel-Aty, A. Nevarez and J. Santos. “Investigation of Safety Influence Area for Four-Leg Signalized Intersections- Nationwide Survey and Empirical Inquiry.” Presented at the Transportation Research Board 87th Annual Meeting, Washington, DC, 2008.

Wilmot, C. and M. Khanal. “Effects of speed limits on speed and safety: a review.” Transport Reviews, Vol. 19, No. 4, 315-329, 1999.

Yau, K. K. W. “Risk factors affecting the severity of single vehicle traffic accidents in Hong Kong.” Accident Analysis & Prevention, Volume 36, No. 3, 333-340, 2004.