Fall 1-31-2013

Driver injury severity at u.s. highway-rail crossings

Wei Hao
New Jersey Institute of Technology

ABSTRACT

DRIVER INJURY SEVERITY AT U.S. HIGHWAY-RAIL CROSSINGS

by
Wei Hao

There are approximately 240,000 highway-rail grade crossings in the United States and highway-rail grade crossing areas have been considered in this study as these are locations where crashes frequently occur. Existing studies on crash models at highway-rail grade crossings can be classified into two categories: accident frequency prediction models and driver injury severity models. Accident frequency prediction at highway-rail grade crossings have been investigated by previous studies using varied statistical models. Few studies, however, have focused on driver injury severity studies. Three drawbacks will be addressed in this research including limitations in traditional highway-rail grade crossings studies, limited models to study driver injury severity, and the relatively small databases. Three driver injury severity models are developed including overall model, driver injury severity model with respect to control devices, and driver injury severity model with respect to age and gender. Based on the model study, it is found that older drivers are more susceptible than younger drivers to cause an increase in severity, an increase in severity under bad weather condition, and improving highway pavement will significantly reduce driver injury severity at passive control highway-rail grade crossings, etc.

DRIVER INJURY SEVERITY AT U.S. HIGHWAY-RAIL CROSSINGS

by
Wei Hao

A Dissertation
Submitted to the Faculty of
New Jersey Institute of Technology
in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy in Civil Engineering

Department of Civil and Environmental Engineering

January 2013

.

APPROVAL PAGE

DRIVER INJURY SEVERITY AT U.S. HIGHWAY-RAIL CROSSINGS

Wei Hao

Dr. Janice Daniel, Dissertation Advisor — Date
Associate Professor of Civil and Environmental Engineering, NJIT

Dr. Taha Marhaba, Committee Member — Date
Professor of Civil and Environmental Engineering, NJIT

Dr. Jian Yang, Committee Member — Date
Associate Professor of Business School, Rutgers University

Dr. Athanassios Bladikas, Committee Member — Date
Associate Professor of Mechanical and Industrial Engineering, NJIT

Dr. Wenge Guo, Committee Member — Date
Assistant Professor of Mathematical Sciences, NJIT

BIOGRAPHICAL SKETCH

Author: Wei Hao

Degree: Doctor of Philosophy

Date: January 2013

Undergraduate and Graduate Education:

- Doctor of Philosophy in Civil and Environmental Engineering,
  New Jersey Institute of Technology, Newark, NJ, 2013
- Bachelor of Science in Civil Engineering,
  China University of Mining and Technology, Xuzhou, Jiangsu, P. R. China, 2002

Major: Transportation Engineering

Presentations and Publications:

Hao,W. and Janice D. (2012). "Driver Injury Severity Study at Highway-rail Grade Crossings in the United." Accepted by *Transportation Research Board Annual Meeting in The Year 2013*. Washington D.C.

Liu, R. and Wei H. (2012). "Bring Economic Benefit to Disadvantaged Populations." Presented on *Transportation Research Board Annual Meeting in The Year 2012*. Washington D.C.

Liu,R. and Wei H. (2011). "Potential Issues and Challenges in Implementing the Railroad Safety Improvement Act of 2008 Based on a National Survey." *Journal of Transportation Safety & Security*, Vol. 3, No. 4, pp. 252-271.

Hao,W. and Mingming C. (2005). "Analysis on Shrinkage Split of Concrete." *Shan Xi Architecture of Journal*. Vol. 31, No. 13, pp. 102-103.

Cao,M. and Wei H. (2005). "Experimental Research on Deformation Law of Soil under Blast-enlargement Loads." *Shan Xi Architecture Journal*, Vol. 31, No. 13, pp. 50-51.

To my beloved God, for his every second with me and for giving me wisdom and guidance throughout my life.

ACKNOWLEDGMENT

I would like to express my deep and sincere gratitude to Dr. Janice Daniel, who not only served as my dissertation advisor, but also provided valuable and countless resources, guidance and encouragement. I am so fortunate to have this kind and knowledgeable advisor who helped me overcome many obstacles and guided me to a new area of interest.

Special thanks are given to Dr. Taha Marhaba, Dr. Jian Yang, Dr. Athanassios Bladikas, and Dr. Wenge Guo for actively participating in my committee and for giving many helpful comments.

Last, but not least, I would like to thank my family. I especially thank my mother, Zhuanling Wang and my father, Budeng Hao, for their encouragement and support in my life.

TABLE OF CONTENTS

TABLE OF CONTENTS
(Continued)

TABLE OF CONTENTS
(Continued)

Chapter **Page**

LIST OF TABLES

LIST OF TABLES
(Continued)

Table **Page**

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

From 1980 to 2010, the number of grade crossing collisions between trains and highway-users fell by 81 percent; corresponding injuries fell by 79 percent; and associated fatalities fell by 69 percent. The varied efforts to improve safety yielded positive results. Although there has been a reduction in the number of collisions, this number is still high and needs to be further reduced (AAR, 2011).

There are more than 250,000 highway-railway grade crossings in the U.S covering a wide range of physical characteristics, control devices and usage. On average, a pedestrian or a vehicle is hit by a train every two hours in the United States. Among all rail-related fatalities, 90% are connected with grade crossing and trespassing incidents (FRA, 2011).

1.1 Problem Statement

Although many studies have been performed to reducing railway highway grade crossing accidents, there are still critical drawbacks in the existing research. Three drawbacks will be addressed in this research including limitations in the modeling approach used in traditional highway-rail grade crossings studies, in the ability of models to predict crashes by control type, and the use of relatively small databases in the model development.

The critical drawback is the limitation in the types of research on traditional highway-rail grade crossings studies. Conventional highway-rail grade crossings studies consider accident prediction models to estimate the frequency of accidents occurring at highway-railway crossings. However, few research studies identify the factors of crashes associated with driver injury severity.

The second limitation is the limited models used to study driver injury severity at highway-rail grade crossings. In previous studies, McCollister (2007) and Hu (2009) both used logit model to investigate key factors for accident severity at railroad grade crossings. However, the inherent ordered relationship of the accident, injury, and fatality was not included. The ideal model should consider the accident, injury and fatality data together using an ordered model. This research will use the following definitions: "Property Damage Only" represents only collisions between vehicle and train; "Injury" is a body wound or shock produced by an accident; "fatality" means death caused by an accident.

The third limitation in highway-rail safety modeling is the limit data sources used by previous studies. Many of the models developed for driver injury severity at highway-rail grade crossings have been developed using datasets for partial area data. For example, Austin (2002) provided an alternative accident prediction model for rail-highway crossings comprising a six-state sample for a 2-year time period. The selected states included California, Montana, Texas, Illinois, Georgia and New York with a sample of 80,962 highway-rail crossings. Researchers considered a total of 1538 highway-rail crossing accidents occurring from January 1997 to December 1998.

In this research, the measure taken to address these limitations includes the use of data from the FRA (Federal Railroad Administration). This data has several advantages including that: 1. It includes all United States' highway-rail crossings; 2. A comprehensive list of variables is provided including transit-control devices, highway and vehicle characteristics, railway and train characteristics, human factors and environmental factors; 3. Driver's injury severity is treated from ordered aspect, meaning the injury levels are ordered from no-injury to the highest injury level, 0 (Property Damage Only), 1 (injured), and 2 (Fatality).

1.2 Research Objective

This study aims to develop driver injury severity models for highway rail-grade crossings using FRA data. Significant factors that have the greatest impact of highway-rail crossing will be identified. The specific objectives are to:

1. To develop a highway railroad grade crossing injury model for all drivers, an overall model, using an ordered probit model and to identify the factors that would influence the injury severity. The factors will be developed by identifying the relationship between the injury severity and a set of independent variables.

2. To develop a driver injury severity model with respect to varied type of control devices (passive control and active control) in order to understand the characteristic of driver injury severity under different control devices.

3. To develop a driver injury severity model with respect to driver's age and gender in order to understand the characteristic of driver injury severity for different driver's group.

1.3 Dissertation Organization

In our study, this dissertation will be organized into six chapters.

Chapter 1 presents the research problem statement and objectives.

Chapter 2 gives the overall literature review of current studies related to railroad highway grade crossing safety.

Chapter 3 describes the research methodologies.

Chapter 4 gives the data processing section.

Chapter 5 provides the research model results and analysis.

Chapter 6 summarizes the conclusions and future works.

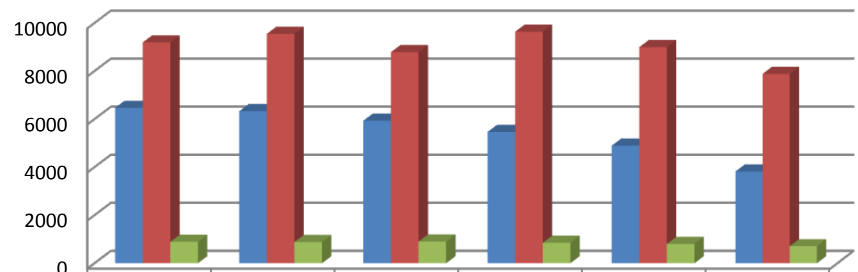
CHAPTER 2

LITERATURE REVIEW

The intent of the literature review is to identify previous research on highway-rail grade crossing studies and past studies dealing with driver injury severity. The literature review is divided into five sections. The intent of Section 2.1 is to investigate the importance of highway-rail grade crossing safety studies. The Section 2.2 aims to find the factors influencing highway-rail grade crossing safety. Section 2.3 deals with current studies of highway-rail grade crossing. Previous highway-rail grade-crossing studies have been conducted to analyze the collision frequency. However, few studies have been conducted on driver injury severity compared to collision frequency studies. The fourth Section 2.4 discusses model building on driver injury severity. Section 2.5 develops specifications of driver injury severity studies from two aspects. The first specification is from the control device aspect and the second is looking at driver’s age and gender.

2.1 Importance of Highway-Rail Grade Crossing Study

Rail transit is considered one of the safest modes of transportation. Every weekday there are more than 7 million people who board rail transit vehicles in the United States (Peterman, 2009). Over the past several decades, great strides have been made in reducing the number of highway railroad grade crossing collisions due to the efforts of federal, state and local governments; railroads; and through organizations such as Operation lifesaver Inc, a nationwide, non-profit public information program to reduce collisions, injuries and fatalities at highway-rail crossings. With nearly a quarter of a million railroad and highway crossings in the U.S., improving grade crossing safety is an enormous challenge that takes the combined efforts of railroads, public safety officials, and the general public (Ries, 2007). An examination of the Figure 2.1 shows that railroad accidents have decreased by 40% from 6470 in 2004 to 3818 in 2009.



	2004	2005	2006	2007	2008	2009
Accidents	6470	6332	5936	5466	4895	3818
Injuries	9194	9550	8790	9638	8996	7882
Fatalities	891	884	903	851	804	708

Figure 2.1 Railroad accidents, injuries, and fatalities from 2004 to 2009.
(Source: Bureau of Transportation Statistics, 2011)

There are approximately 240,000 highway-rail grade crossings in the United States. Among these crossings, around 39 percent are private highway-rail grade crossings and the remainder, or 61 percent, are public highway-rail grade crossings (FRA, 2010). Between 2000 and 2010, there has been a reduction in the number of incidents at highway-rail grade crossings from 3502 (2000) to 2017 (2010). At the same time, the number of fatalities at highway-rail grade crossings also reduced from 425 (2000) to 256 (2010) (See Figure 2.2).

Over the past several decades, great strides have been made in reducing the numbers of railroad highway grade crossing collisions due to the efforts of federal, state and local governments; railroads; and through organizations such as Operation lifesaver

Inc, a nationwide, non-profit public information program to reduce collisions, injuries and fatalities at highway-rail crossings. With nearly a quarter of a million railroad and highway crossings in the U.S., improving grade crossing safety is an enormous challenge that takes the combined efforts of railroads, public safety officials, and the general public (Ries, 2007).

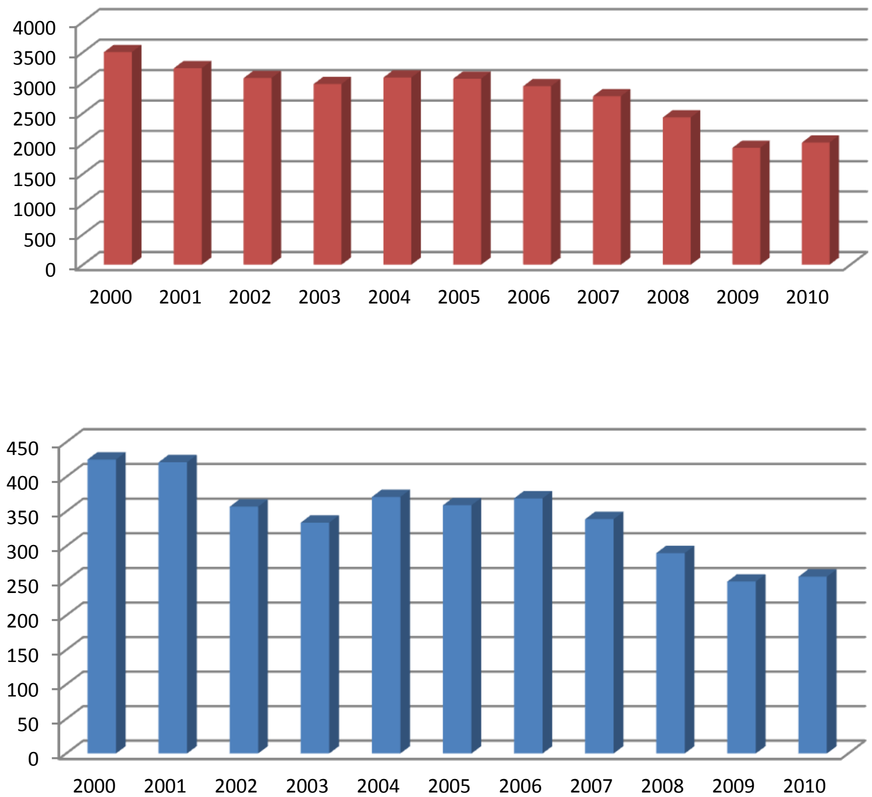


Figure 2.2 Highway-rail grade crossings information.

(Source: Federal Railroad Administration Office of Safety Analysis, 2012)

2.2 Factors Influencing Rail-Highway Grade Crossing Safety

Several factors affect the safety of highway-rail grade crossings. Control devices, human factors, vehicle factors, schedule factors, environment factors and education factors are several of these factors and are described in this literature review. The literature provides a discussion of control devices, such as signs, pavement markings, flashing lights, and automatic gates. In addition, highway factors and railway factors will be discussed in detail in Subsection 2.3.1.3 “Previous Practices of Highway-Rail Crossing Crash Models”.

2.2.1 Control Devices

The type of warning device used at a highway rail-grade crossing has a significant effect on the risk at grade crossings (Farr, 1987). There are two types of warning devices: passive and active. Passive traffic control devices give static information of warning, guidance, and mandatory action for the driver. Passive traffic control systems consist of signs, pavement markings, and grade crossing illumination. Passive crossings lack train-activated warning devices and display signs and pavement markings to identify the location of the crossing and to direct the attention of the motorist, bicyclist, or pedestrian. Active traffic control systems include flashing signals, bells and automatic gates. Active traffic control devices are those that give warning to the approach or presence of a train. Active control devices are supplemented with the same signs and pavement markings used for passive control. In sum, active crossings contain devices that warn drivers of the approach or presence of a train. Noyce (1998) studied enhancements to traffic control devices at passive highway-railroad grade crossings. The objective of the research was to test and evaluate an improved method for communicating with drivers at passive highway-rail grade crossings. The enhanced sign system involved a full-sized strobe light, a shield, and a loop detector. Power for the loop detector and strobe light was provided by a solar charged 12-volt battery. An effective method to determine if the system improved safety at passive crossings was to evaluate the crash rates at the crossing before and after installation. This research found that the enhanced sign system was effective in reducing speeds and attracting drivers' attention to highway-railroad grade crossings.

Peck (2010) studied the differences in the United States at public and private highway-rail crossings. Figure 2.3 shows the number of incidents by warning decides at private and public crossings. The highest number of incidents at private crossings occurs at locations with passive control that are equipped with cross bucks or stop signs. The highest number of incidents at public crossings occurs at crossings equipped with cross bucks, but a high percentage has flashing lights or flashing lights and gates.

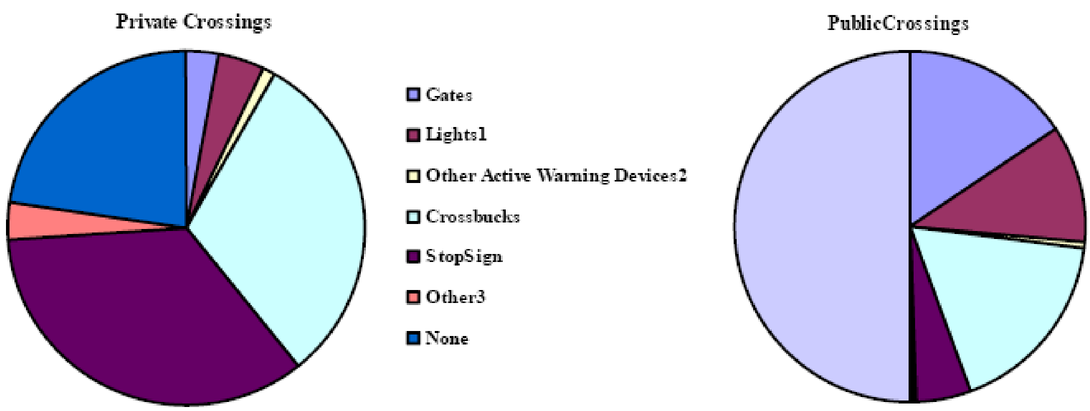


Figure 2.3 Numbers of incidents by warning device.
(Source: Peck et al., 2011)

2.2.2 Human Factors

Rahimi (2001) conducted research to explore the hypothesis that driver decision-making styles influence high-way-rail crossing accidents. From his study, one-third of rail accidents and over 80 percent of train collisions are caused by human error. In this research, a "descriptive-differential" approach was used to match the driver's decision style, driving task demands, and then determine the fit to the environmental factors of highway-rail crossing. An analysis of variance experiment was designed with three independent variables including “driver decision style”, “driver time pressure” and “intersection complexity”. The decision style modes included in this research were: 1) the manner in which the driver reacts to a given crossing situation; and 2) the manner of interaction with other environment factors including time pressures and mental load. The research concludes that decision styles are important factors to understanding HRC driving activities. This research could provide insights into experimental design approach and help us understand human factor as a significant factor to influence highway-rail crossing safety. However, this research is lacking a real data source to validate their conclusions and we will use FRA data to prove human factor as a key factor in our research.

A study plan by Anandarao and Martland (1998) provides the application of probabilistic risk assessment techniques to determine the efficacy of the various level crossing safety devices in Japan. An exploratory analysis method to determine the factors affecting the risk of a level crossing accident is provided. The methodology uses two questions to determine the safety of a transportation system: 1) what will happen? And 2) what will be acceptable? The first question investigates risk analysis using techniques from engineering and probability theory. The second question involves value judgments on the part of risk assessment study. The study states that the most important level crossing attributes affecting highway-rail crossing accidents could be summarized as: rail traffic volume, road traffic volume, location of the crossing, visibility of the crossing, road gradient, distance to the nearest road intersection, number of tracks, and the type of safety devices. The following gives the detailed factors affecting the risk of a level crossing accident: (1) Crossings with visibility less than 20 m cause a 50% higher accident rate than crossings with visibility greater than 20m; (2) the accident rate proportionately increase as the number of tracks increases; (3) Crossings with low rail and road traffic volume are riskier per train than crossings with high rail and road traffic volumes. At low rail and road traffic volume crossings, the possibility a vehicle will go through the crossing with the warning bell is ringing is high since there is no vehicle in front of it and the risk increases if the rail traffic is low since the vehicle might not be aware of the approach of a train. The results of this crossing safety study showed that the leading cause of the crossing accidents was the driver's non-compliance of traffic control devices or to simply ignore all warnings.

Jonsen (2007) presented studies relevant to human factors and effects of safety measures on passive railroad-highway grade crossings. The purpose of that study was to describe users' judgment of speed and distance related to trains at passive railroad-highway grade crossings. The contribution of this paper was a complete literature review of studies on human factors including judgment of speed and distance, base critical lag and clearance time for the road traffic, and sight distances at highway-rail grade crossings. The limitation is that the paper lacks the use of a full dataset to prove their ideas. Road users' perceptual underestimation of trains' time-to-arrival at grade crossings become larger with the closer trains at crossings due to systematic illusions within the human vision. In addition, creating a perpendicular crossing, reducing gradients, and increasing sight distance could make railroad grade crossings safer. Furthermore, an educational campaign could also improve safety at grade crossings.

2.2.3 Area Type

McCollister (2007)'s probability model, which will be discussed later in Section 2.3.3, also considers the area type as variables in a model to predict the probability of accidents, injuries and fatalities at highway-railway crossings. From the result, the presence of commercial areas is associated with higher accidents. The commercial area is correlated with relatively more complicated traffic activities and drivers may be unfamiliar with the crossing. The contribution of this paper is that area type should be included as a variable in the logit model. However, the paper lacks an explanation of residential and industrial areas. In the proposed dissertation, all three area types will be considered into our model analysis.

2.2.4 Education and Law Enforcement

Sposato (2006) provided studies on the impact of public education and enforcement on driver and pedestrian behavior at highway-rail grade crossings. The purpose of the study was to determine whether community education and/or enforcement activities were successful in significantly reducing the violation rate at highway-rail grade crossings. To evaluate the effectiveness of education programs, researchers measured the number of the motor vehicle and pedestrian violations occurring before, during, and after the Public Education and Enforcement Research Study (PEERS) program. The evaluation team used video cameras to observe the frequency with which motorists and pedestrians violated the traffic control devices. The PEERS program observes driver and pedestrian behavior at highway-rail grade crossings before and after the program was implemented. The program was considered successful if the violation rate was reduced by 50 percent after the program was implemented. This research provides significant and meaningful results about the effectiveness of education and enforcement activities. Two conclusions were made in this research: (1) crossing demographics and characteristics were determined to play an important role in the study; and (2) Community education and/or enforcement activities were successful in significantly reducing the violation rate at highway-rail grade crossings. The enhanced education and enforcement activities could be evaluated by using a cost benefit study. The cost benefit ratio is estimated as the cost of law enforcement versus the potential lives saved

Savage (2005) conducted public education to improve rail-highway crossing safety. The public education program talked in this paper is called operation lifesaver (OL). Operation lifesaver programs were established in each state to promote education and awareness of railroad related hazards, especially the need to appreciate the risks when traversing grade crossings. This paper uses a negative binomial regression to estimate the impact of Operation Lifesaver activity across states and from year-to-year in individual states will be related to the number of collisions and fatalities at highway-rail crossings. The data set consists of a collection of 46 states for the years from 1996 to 2002. Dependent variable is number of incidents in a state in a given year at public crossings and explanatory variables include the levels of rail and highway traffic (AADT, trains per day, etc.), the warning devices factors, and highway safety performance variable. The analysis finds that increasing the amount of educational activity will reduce the number of collisions; however the effect on the number of deaths could not be concluded with statistical certainty.

2.3 Current Highway-Rail Crossing Studies

Highway-rail crossing accident injury and fatality rates are much higher than other types of traffic collision due to the significant mass difference between traffic and train. As a result, compared to highway intersections, highway-rail grade crossings should be paid more attention for collision modeling and prediction analysis. However, there are few studies conducted on highway-rail crossing studies compared with highway intersection studies. Table 2.1 lists thirteen studies of highway-rail crash conducted in a time ranging from the late 90s to 2011. These studies can be classified as two types: collision frequency study and collision injury study.

A number of previous highway-rail grade-crossing studies have been conducted to analyze the collision frequency. These studies will be discussed in Section 2.3.1. However, few studies have been developed to analyze the vehicle driver's injury in their highway-rail grade crossing studies which will be discussed in Section 2.3.2.

2.3.1 Previous Highway-Rail Crossing Collision Frequency Study

Over the last few years, a large number of collision frequency models have been developed. Traditional accident prediction models could be classified as two types: absolute and relative risk models. Absolute models estimate the "expected number of collisions" at a given crossing for a given period. Relative risk models estimate a "hazard index" representing the relative risk of one crossing compared to another. In addition, statistical models including the Poisson, Negative Binomial and discrete choice models are also developed recently to analyze factors influencing collision frequency.

suited for real-data applications. In recent years new adaptive algorithms have been suggested for subspace tracking (Patel, 1998), (Valdez, 1999).

Table 2.1 Previous Study of Highway-Rail Crossing Studies

Author	Yea	Paper Title
Gitelamn *et al*.	1997	The evaluation of road-rail crossing safety with limited accident statistics
Austin *et al.*	2002	An alternative accident prediction model for highway-rail interfaces
Saccomanno *et al.*	2004	Risk-based model for identifying highway-rail grade crossing blackspots
Miranda-moreno *et al.*	2005	Alternative risk models for ranking locations for safety improvement
Oh *et al.*	2006	Accident prediction model for railway-highway interfaces
McCollister *et al.*	2007	A model to predict the probability of highway rail crossing accidents
Saccomanno *et al.*	2007	Estimating countermeasure effects for reducing collisions at highway-railway grade crossings
Park *et al.*	2007	Reducing treatment selection bias for estimating treatment effects using propensity score method
Miranda *et al.*	2009	How to incorporate accident severity and vehicle occupancy into the hotspot identification process?
Raub *et al.*	2009	Examination of Highway-Rail Grade Crossing Collisions Nationally from 1998 to 2007
Hu *et al.*	2010	Investigation of key factors for accident severity at railroad grade crossings
Yan *et al.*	2011	Using hierarchical tree-based regression model to predict train-vehicle crashes at passive highway-rail grade crossings
Eluru *et al.*	2012	A latent class modeling approach for identifying vehicle driver injury severity factors at highway-

The Peabody Dimmick Formula was developed in 1941 using accident data from rural railway-highway crossings in 29 states in US. The model estimates the expected number of accidents in the highway-rail grade crossing in 5 years using four parameters including average annual daily traffic (AADT), the average daily train traffic (T), protection coefficient indicative of warning devices (P) and additional parameter (K).

Table 2.2 Typical Absolute Model Studies

Peabody Dimmick Formula	USDOT Accident Prediction Model
$A_5 = \frac{1.28\left(V^{0.17} * T^{0.151}\right)}{P^{0.0171} + K}$	$a = K * EI * DT * MS * HP * HL * HT$
A_5 = the expected number of accidents in 5 years V = average annual daily traffic (AADT) T = average daily train traffic P = protection coefficient indicative of warning device presents K = the additional parameters	a = un-normalized initial crash prediction, in crashes per year at the crossing K = formula constant EI = factor for exposure index based on product of highway and train traffic DT = factor for number of through trains per day during daylight MS = factor for maximum timetable speed MT = factor for number of main tracks HP = factor for highway paved (yes or no) HL = factor for number of highway lanes
The expected number of accidents in 5 years	A formula containing geometric and traffic factors from the inventory file A formula involving crash history A formula incorporating the effect of the existing warning devices

The US-DOT model which was developed in the 1980s is the typical absolute model and recognized as the industry standard for collision risk prediction at highway-railway grade crossings. Compared to the Peabody Dimmick Formula, US-DOT Formula has additional factors including the exposure index which is based on the product of highway and train traffic, number of through trains per day during daylight, a factor for maximum timetable speed, a factor for number of main tracks, a factor for whether the highway is paved (yes or no) and a factor for number of highway lanes.

The next step in highway-rail crossing accident prediction method was the New Hampshire Index, California's Hazard Rating Formula and Connecticut's Hazard Rating Formula. These methods differ by states and the formulae are provided in Table 2.3. The New Hampshire index uses three factors: number of vehicles per day, number of trains per day and a protection factor based on the type of crossing. California's Hazard Rating Formula uses four variables: number of vehicles, number of trains, crossing protection type and the crash history. Connecticut's Hazard Rating Formula is similar to California Rating Formula except it uses a ten-year crash history while California uses a five-year history. Several studies were conducted over the last few decades using different types of road collision models. The Poisson regression is usually a good modeling start due to crash data with approximately Poisson distribution. When data are observed with over dispersion, some modifications to the standard Poisson regression are available. The most common variations include the negative binomial model and zero-inflated negative binomial models. In addition, the less common model is the gamma probability count model.

Table 2.3 Hazard Index Formula Studies

New Hampshire Index:	**California's Hazard Rating Formula**	**Connecticut's Hazard Rating Formula**
$HI = V * T * PF$	$HI = \frac{V * T * PF}{1000} + H$	$HI = \frac{(T+1)(A+1) * AADT * PF}{100}$
V =the average annual daily traffic (AADT) T = average daily train traffic PF = the protection factor indicative of warning device present	V = number of vehicles T = number of trains PF = protection factor form H = crash history= total number of crashes within the last ten years *3	T = trains movements per day A = number of vehicle/ train crashes in last 5 years $AADT$ = Annual Average Daily Traffic PF= Protection Factor
The protection factor varies from state to state and accurately predicting railway-highway crossing accidents	Does not compute the number of crashes but rather produces a hazard index as an alternative for the number of crashes The crossing with the highest calculated index	Only difference is the crash history period with a ten-year crash history in Connecticut compared with five-year history in California.

The main purpose of previous research using Poisson and Negative Binomial models was to establish statistical relationships between collisions and various road geometry and traffic attributes. The following describes various types of highway-rail crossing safety models.

In the Poisson regression model (Oh, 2006), the expected number of crashes follows a Poisson distribution, the expected crash count for the ith crossing is given as $\hat{y}_i$, i=1,…,N, is a function of covariates X_{ij} , i=1,…,N, j= 1,…,M,

$$y_i \sim Poi(\lambda_i);\hat{\lambda}_i = \exp(\beta_0 X_{i0} + \beta_1 X_{i1} + ... + \beta_M X_{iM}) = \exp\left(\sum_{j=1}^{M} \beta_j X_{ij}\right) \quad (2.1)$$

Where the β_j's are the estimated regression coefficients across covariates j= 1,…,M (for the slope intercept model the first covariate is a vector of 1's) averaged across crossings i=1,…,N. Because the Poisson regression model is heteroscedastic, the model coefficients are estimated by maximum likelihood methods. The likelihood function is given as:

$$L(\beta) = \prod_i \frac{\exp[-\exp(\beta X_i)][\exp(\beta X_i)]^{y_i}}{y_i!} \quad (2.2)$$

The maximum possible value of the likelihood set occurs if the model fits the data exactly, resulting in a value of 0 for the likelihood function. In addition, if the mean of the crash counts is not equal to the variance, the data is said to be over dispersed.

The negative binomial model (Oh, 2006) takes the relationship between the expected number of accidents and the M parameters, $X_{i1,} X_{i2,} ..., X_{im}$

$$y_i \sim Poi(\lambda_i);\hat{\lambda}_i = \exp(\beta_0 X_{i0} + \beta_1 X_{i1} + ... + \beta_M X_{iM} + \varepsilon_i) = \exp\left(\sum_{j=1}^{M} \beta_j X_{ij} + \varepsilon_i\right) \quad (2.3)$$

Where $\exp\left(\varepsilon_i\right)$ is distributed as gamma with mean 1and variance α^2 .

The characteristics of the negative binomial model is listed as

1) The effect of the error term in the negative binomial regression model allows for over dispersion of the variance.

$$Var\left(y_i\right)=E\left(y_i\right)+\alpha E\left(y_i\right)^2 \qquad (2.4)$$

Where α is the over dispersion parameter.

2) If the over dispersion parameter, α , equals 0, the negative binomial reduces to the Poisson model. The larger the value of α , the more variability there is associated with the mean $\hat{\lambda}_i$. The coefficients β_j are estimated by maximizing the log likelihood $\log_e L\left(\beta\right)$.

The gamma model was proposed and discussed by Oh (2006) .The gamma probability model is given as:

$$\Pr\left[y_i = j\right]=Gam\left(\alpha j,\lambda_i\right)-Gam\left(\alpha j+\alpha,\lambda_i\right) \qquad (2.5)$$

Where:

$$\lambda_i = \exp\left(\beta' X_i\right) \qquad (2.6)$$

$$Gam\left(\alpha j,\lambda_i\right)=1 \text{ if j=0, or } \frac{1}{\Gamma\left(\alpha j\right)}\int_0^{\lambda_i} \mu^{\alpha j-1}e^{-\mu}d\mu \qquad (2.7)$$

If $j>0,\ j=0,1,\ldots$

The gamma model is used when the crash mean is greater than the crash variance.

The dispersion parameter is again α in three scenarios: 1. under dispersion if $\alpha > 1$;2. Over dispersion if $\alpha < 1$; 3.Equidispersion if $\alpha = 1$, which reduces the gamma probability model to the Poisson model.

Articles by Oh (2006) and Austin (2002) are referenced because they systematically discussed traditional accident prediction models including Poisson and Negative Binomial regression models. In addition, research performed by Saccomanno (2004) is provided due to the detailed discussion of “highway, railway, and vehicle” factors to influence highway-rail crossing accident.

Oh (2006) developed an accident prediction to examine factors connected with railroad crossing crashes. In this paper, the author conducted the literature review on traditional highway-rail crossing collision frequency mode including Peabody Dimmick Formula, US DOT formula, and New Hampshire Index. After that, the gamma probability model statistical model was given to examine the relationships between crossing accidents and features of crossings. The gamma probability model is a flexible model and is relatively new in transportation safety research. This paper uses highway-rail grade crossing data from Korea where there were 402 accidents between 1998 and 2002. This paper not only gives insights from a model aspect but provides interesting research variables including daily traffic volume, daily train volumes, proximity of commercial area, distance of train detector from crossing, time duration between the activation of warning signals and the activation of gates, and the presence of speed hump. In addition, number of tracks and average daily railway traffic (trains per day) is listed as railway characteristics to analyze the train influencing factors. This research suggests that more studies should include examination of driver warning devices, such as devices which detect and warn approaching vehicles, trains, or both.

Saccomanno (2004) presented a risk-based model to identify highway-rail grade crossing blackspots using Canada grade crossing data over the last 20 years. The following section listed key variables used in the Poisson regression models: track angle, number of tracks, train speed, road speed, surface width, road class, highway paved, warning type, AADT, number of trains daily, and number of collisions. According to Saccomanno's research, traffic exposure (log of cross product of AADT and the number of daily trains) were found to be the most important factor for expected frequency of collisions at highway-rail grade crossings. The findings of this research are that crash frequency is dependent on types of warning device. For passive crossings (signs only), train speed was found to be the highest explanation for the expected frequency of collisions per year. For active crossings with flashing lights, the significant factors were train speed and road surface. For crossings with gates, road speed and number of tracks were found to be highest prediction factors. As a result, the risk models developed in this research explain that fewer collisions occur at crossings equipped with flashing lights and gates than at crossings with signs.

The research objective by Austin (2002) was to identify an alternative accident prediction model for Rail-highway crossings using negative binomial regression. The data sample for this investigation included a wide geographical coverage of a six-state sample for a 2-year time period. The selected states included California, Montana, Texas, Illinois, Georgia and New York with a sample of 80,962 highway-rail crossings. Researchers considered a total of 1538 highway-rail crossing accidents occurring from January 1997 to December 1998. Traffic characteristics, roadway characteristics, and crossing characteristics were considered in this research. For the roadway characteristics, four elements of the highway were included: roadway type, surface width, traffic volume and control devices. If a highway is paved, there is a higher likelihood of an accident than if it is gravel. Second, surface width, which is the measured distance of the highway at the crossing approach. The surface width could be taken as the number of lanes. The greater the number of traffic lanes, the higher the highway-rail crossing collision frequency. Third, the higher the traffic volume on the highway, the larger number of vehicles that are exposed to conflicts with train movements and the greater the probability of collision. In addition, the presence of gates and highway traffic signals were found to significantly reduce crossing accident frequency. On the contrary, the presence of stop signs, flashing lights, and bells were found to increase predicted collision. In sum, the author has considered traffic, roadway and crossing characteristics to develop an alternative highway-rail crossing accident prediction model.

2.3.2 Injury Severity Study in Highway-Rail Crossing Study

Eluru (2012) developed a latent class model to identify vehicle driver injury severity factors at highway-railway crossings. The traditional ordered response model assumed that the effect of various factors on injury severity to be constant across all accidents. The latent model applied an innovative latent segmentation model addressing the issues to evaluate the effects of various factors on injury severity at highway-railway grade crossings. The dataset is from U.S. Federal Railroad Administration database highway-rail grade crossing inventory and collision data including 14532 crossings from 1997 to 2006. The factors which found to be significant influencing injury severity included driver age, time of accident, presence of snow/ or rain, vehicle role in the crash and motorist action. However, the author just included the public grade crossings on the main railway line and collision involving passenger vehicles. In reality, accident happening in private crossings and commercial vehicles should be considered in the further studies.

Miranda-Moreno (2009) modeled and estimated the severity levels of each individual involved in an accident using a multinomial model. A sample of highway-railway intersections in Canada comprising 1773 crossings is considered in the research case. The collision database for the period from 1997 to2004 with 941 highway-railway grade crossing collisions was included. Specially, the author considered the total risk as the product of accident frequency and expected consequence. However, this research limited to provide only trains speed and posted speed limit variables in their analysis and neglected to provide many other potential exogenous variables.

Hu (2009) conducted a logit model to investigate key factors for accident severity at railroad grade crossings. The dataset is from the railway police and the Taiwan Rail Administration (TRA) at railroad grade crossings in Taiwan which collected from 1995 to 1997. The original dataset included railway features, highway features, crossing features, traffic control, others. It was found that variables such as the number of daily trains, number of daily trucks, obstacle detection devices had positive increase in severity accidents. However, traffic control devices and management tools are surprised to find not significant to cause an increase in severity. The limitation of this research located in the few talking about the driver demographics studies.

McCollister (2007) developed an injury severity model to predict the probability of accidents, injuries and fatalities at highway-railway crossings. A logistic regression method was adopted as the methodology for estimating the probability of fatality for vehicular occupants and two databases from Federal Railroad Administration (FRA) were used in estimating the injury severity model. Train speed, number of trains, percent of truck, traffic/lanes, angle, traffic control devices, area type, and accident history are used as the research variables in estimating crash injury. After model testing, train speed, number of trains, percent of trucks, traffic control devices and accident history were found to be significant variables. For the traffic control devices, the coefficients for cross bucks and stop signs are nearly equal. The most significant variables were accident history and traffic congestion. The number of night through trains was very significant, but the number of day through trains was less important. The square root of the maximum speed on a section of track is also highly significant. An interesting result shows that trucks are 60 percent less likely to be involved in a rail-highway crossing crash than a passenger automobile. In addition, more variables should be included, such as driver's information (age and gender), and weather conditions (visibility and clarity). The contribution of study is the author separately used logit models to consider the accident, injury, and fatality data. However, the inherent ordered relationship of the accident, injury, and fatality was not included. The ideal model should consider the accident, injury and fatality data together using an ordered model (Accident: collisions between vehicle and train; Injury: a body wound or shock produced by accident; fatality: death caused by accident).

2.4 Model Building on Injury Severity Level Study

Based on the previously discussed review, although several studies have been conducted to investigate highway-rail grade crossings, there is little research on modeling injury severity studies at highway-rail grade crossings. For this reason, the highway injury severity literature would be the best source to guide the development of highway-rail grade crossing injury severity modeling. Lots of methodological methods have been conducted to analyze highway crash severity. The dependent variable is the key factor to determine the frame of the methodology. The dependent variable of current existing crash severity models could be either a binary response (e.g. injury or nonjury) or a multiple response (.e.g. fatality, injury, or noninjury). The multiple responses of the dependent variables could be classified as ordered or unordered. Five types of models including Binary model, multinomial logit, nested logit model, ordered probit model, and mixed model are presented here to make up the literature section.

2.4.1 Binary Model

Huang (2008) conducted Bayesian hierarchical logit model to identify the significant factors influencing the severity of driver injury and vehicle damage in traffic crashes. In this study, crash data is collected from Singapore from 2003 to 2005. There are total 19832 reported crashes in this period and 4095 occurred at signalized intersection were used in the model. This study provided a way to analyze the potential within-crash correlation study using the hierarchical modeling technique. In detail, the article found that speeding and alcohol use resulted in higher crash severity. The effects of street lighting at night come into play an important role in this study.

Lee (2008) examined the impact of passengers on the driver's crash potential on freeways. A bivariate probit model was developed using the crash record in Orlando, Florida from 1999 to 2003. Based on the bivariate probit model analysis, there are strong correlations between passengers and crash characteristics. Driver's behavior is safer when they are accompanied by passengers and more passengers reduce driver's crash potential. According to this research, younger drivers are strongly recommended to be accompanied by one or more older passengers. In addition, the younger drivers are also recommended to drive slower when they drive with only younger passengers in high speed or low-volume traffic road.

2.4.2 Multinomial Logit Model

Tay (2011) successfully conducted a multinomial logit model to analyze the pedestrian-vehicle crash severity. The purpose of this study was to determine the factors which contribute to the severity of pedestrian-vehicle crashes in South Korea. A number of factors are calibrated to relate crash severity including roadway environment, traffic control devices, weather conditions, pedestrian location, pedestrian and driver's characteristics, pedestrian and vehicle characteristics. The factors identified as increasing the probability of fatal injury included: drivers' sex, age and alcohol intoxication; pedestrians' age and sex; pedestrians' location on crosswalk, intersections, shoulder, freeways; wider roads especially wider than 9 m; vehicle type and size; inclement weather like cloud, fog, snow and rain; time of day such as night time and peak hours; and relatively less urbanized regions. As a result, the fatal and serious crashes were associated with collisions with heavy vehicles; drunk drivers; pedestrians with age over

65 or female, on high speed roads, in bad weather condition, at night. In sum, the findings of this paper could be referenced in our study such as the classification of potential useful variables.

Malyshkina (2009) conducted an application of multinomial logit model to accident-injury severities to capture unobserved heterogeneity in accident data which could relate to detailed weather conditions. The model successfully accounts for the potential of unobserved heterogeneity between two unobserved roadway safeties. The conclusion found was that more roadway safety is correlated with better weather conditions and on the contrary the less frequency is strongly related to adverse weather conditions.

Shankar (1996) explored the use of the multinomial logit model for evaluating injury severities for single-vehicle motorcycle accidents. The research uses 5 years of data from the state of Washington to estimate a multivariate model on motorcycle severity. The influencing factors include environmental factors, roadway conditions, vehicle characteristics, and rider attributes. A number of variables found to influence accident severity suggest a number of important directions for future studies. First, multivehicle accidents should be considered in the further study instead of single-vehicle crash. Second, the dataset here is limited to Washington and more affluent databases are needed for future work.

2.4.3 Nested Logit Model

Savolainen and Mannering (2007) studied motorcyclists' injury severities in single and multi-vehicle crashes using nested logit model. The database used in this paper is from the state of Indiana between January 1, 2003 and October 15, 2005. The important findings present that increasing motorcyclist age is associated with more injuries. In addition, the collision type, roadway characteristics, alcohol consumption, helmet use, unsafe speed play significant roles in crash-injury outcomes.

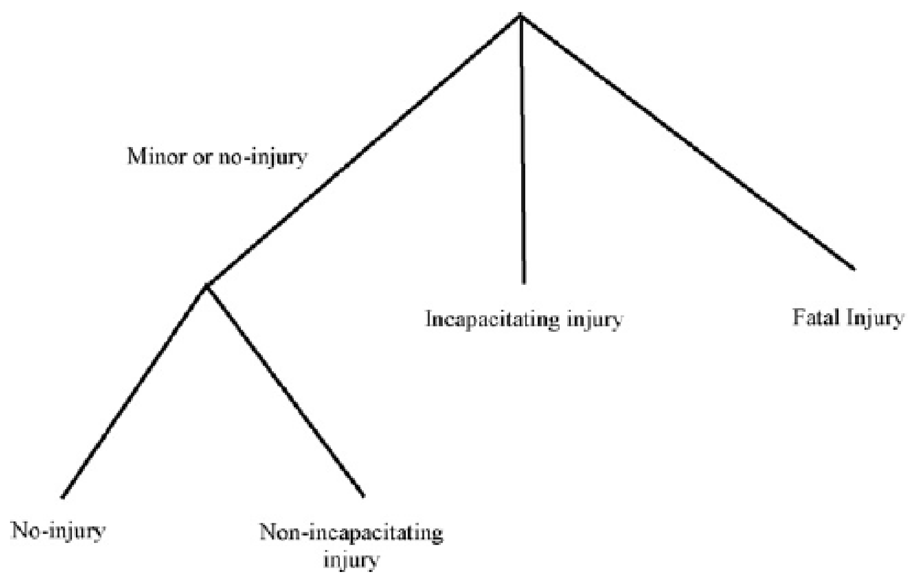


Figure 2.4 Nested logit structure of crash injury severity model.

(Source: Savolainen, P. and F. Mannering, 2007)

2.4.4 Mixed Logit Model

Milton (2008) studied highway accident severities using the mixed logit model. The characteristic of this approach shows that estimated model parameters could vary randomly across roadway segments relating to roadway characteristics, environmental factors, and driver behavior. The findings indicate that volume-related variables such as average daily traffic per lane, average daily truck traffic, truck percentage, and weather conditions are best modeled as the random parameters, while roadway characteristics such as the number of horizontal curves, number of grade breaks per mile and pavement friction are best to be modeled as fixed parameters.

2.4.5 Ordered Probit Model

The following section will provide representative papers using ordered probit models in highway injury studies. These studies could help to understand the type of variables that might be considered in an injury severity model and show how to build the relationship between injury severity and these related variables.

The primary objective of Zhang's (2011) study was to explore the contributing factors influencing the crash injury severity at diverge areas and quantitatively evaluate their impacts. The study uses crash data at selected freeway exit segments in Florida. It is strongly related to our highway-rail grade crossing injury severity studies because it demonstrates the use of the ordered probit model and can also indicate the significant variables which may influence highway-rail crossing safety. It was found that the factors significantly impacting injury severity include number of lanes, speed limits, light condition, weather condition, surrounding land type, alcohol/drug involvement, road surface condition, and shoulder width. The specific finding could be summarized as: 1. One additional lane on mainline will decrease the proportion of no injury crash by 2.1%; 2. Good light and weather condition will increase the probability of no injury by 3.4% and 3.3%, respectively. The alcohol involvement will increase the probability of injured crash by 14.8%. Abdel-Aty (2003) used the ordered probit model to analyze the driver injury severity at multiple locations including roadway sections, signalized intersections, and toll plazas in Central Florida. Factors found to significantly impact the three injury severity models include driver's age, gender, seat belt use, vehicle type, point of impact, and speed ratio. Other factors were specific to the location of the crash. For example, roadway curves and dark lighting conditions contribute to higher probability of injuries on roadway sections. Second, rural areas were found to have a high probability of injuries due to higher speed. Third, driver's errors are found to be a significant variable in the signalized intersections' model. Fourth, if the vehicle is equipped with an electronic toll collection device, there is higher probability that the driver will have an injury related to higher speed in toll plazas. The contribution of this study is to introduce the land use aspect to the driver injury study. It can be used in this dissertation to classify injuries based on different locations and then did model injury severity as a dependent variable from land use aspect correlated with other variables including driver's information, traffic control type, traffic volume and so on.

Kockelman (2001) modeled the driver injury severity to assess risk factors and design issues in roadway travel. The objective of this paper is to examine the risk of different levels sustained under all crash types, two-vehicle crashes, and single-vehicle crashes. The probability of injury severity level is examined by applying an ordered probit regression model recognizing the ordinality of injury level. A variety of factors could come into play when vehicles crash on the road. The study data was derived from the 1998 National Automotive Sampling System General Estimates System (GES) of all police-reported crashes in the U.S. This research concludes that the manner of collision, number of involved vehicles, driver gender, vehicle type, and driver alcohol use would play major roles. The contribution of this study is the separation of different type of vehicle crashes. Based on the findings of this study, pickups and sport utility vehicles are less safe than passenger cars under single-vehicle crash condition. However, these vehicles are safe for the drivers compared with occupants under two-vehicle crashes.

2.5 Driver Injury Severity by Control Device

There are existing literatures on examining the effects various traffic control measures on the accident frequencies. Raub (2006) examined highway-rail grade crossing collisions over 10 years in seven Midwestern states to compare four major classes of warning devices for highway-rail grade crossings. The data covers a 10-year period from 1994 to 2003 for collisions including injuries and fatalities. Several conclusions can be made: 1) gates usually have the lowest collision rates; and 2) collisions at highway-rail crossings with STOP signs are more likely to occur than with other types of warning system. For STOP sign, drivers misjudge the speed of the approaching train and therefore believe they have sufficient time to cross the intersection before the train arrives. Zwahlen and Schnell (2000) compared driver behavior at the standard crossbuck with two experimental reflectorized crossbuck systems in a before-and-after study. The study found that reflectorization increased the time between a noncompliant vehicle crossing the track and the on-coming train. Meeker et al. (1997) provided a comparison of driver behavior at railroad grade crossings with two different protection systems. The effectiveness of a flasher-only protection system was compared with one incorporating flashers and barrier gates for a particular crossing. The addition of the gates significantly reduced the percentage of drivers crossing in front trains from 67% to 38%. Abraham et al. (1998) examined driver behavior at highway-rail grade crossings to determine the difference between gate control and flashers. Drivers tend to commit more violations at the gated highway-rail grade crossings with more traffic control devices compared to crossings with only flashers. The limitation of gated control could be that drivers have better chances of clearing the intersections before the train's arrival in the no-gated control.

There is clear evidence based on the above mentioned studies documenting the decreased risk of train-vehicle collision occurrence as a result of presence of junction control measures. Although exist several studies have already examined the effects of control measures on the highway-rail grade crossing accident frequencies, however no current study was found studying driver injury severity under various control devices at highway-rail grade crossings. As a result, reference studies have been conducted to investiage the injury severity of drivers under various traffic control measures at non-highway rail crossing. Four of these types of studies were selected for review including Haleem (2010), Pai et al (2007), and Zhang et al. (2000). These studies were reviewed because they show information on driver's injury severity varied by different control devices. The recent study performed by Haleem (2010) examined traffic crash injury severity at unsignalized intersections including 2,043 unsignalized intersections in Florida from 2003 to 2006. Based on this study, it was found that higher severity probability is always associated with a reduction of AADT, and an increase of speed limit. In addition, heavily-populated and high-urbanized areas were found to have lower injury severities.

The most related study looking at the relationship between traffic control and injury severity was a study performed by Pai et. al. (2007) that explored the impact of motorcyclist injury severity under various traffic control measures. That study was performed using data from the UK and looked at injury as a function of demographic, vehicle and environmental factors. Although this study did not evaluate highway drivers at highway-rail grade crossings, the results from this research are useful in understanding the impact of traffic control on driver injury at highway-rail grade crossing. The database extracted accident injury from 1999 to 2004 in the UK. Control measures are divided into three categories: 1. Stop, give-way signs or marking; 2. Uncontrolled; 3. Signal measures. The model result suggests that the combined effect of riding in darkness and uncontrolled junction was dangerous to motorcyclists. A reduction of speed limit at unsignalised crossings would be effective to decrease injury severity to allow more reaction time for last-minute braking the moment before impact. Another study by Zhang (2000) investigated factors affecting the severity of motor vehicle traffic crashes involving elderly drivers aged 65 and over between 1988 and 1993 on Ontario public roads. This study indicated that elderly drivers involved in crashes at non-controlled intersections had an increased risk of fatal outcome compared with those involved at controlled intersections.

To sum up, the existing studies have provided valuable insights into the relationship between various factors and driver injury severity. Nevertheless most of these studies focused on collisions happened along roadway segments rather than a specific type of crossings. Without a proper understanding of multiple factors influencing injury levels, the countermeasures based on previous studies could be ineffective. This study attempts to apply appropriate statistical modeling approach to analyze highway-rail crossing data from 2002-2011 in the U.S., exploring the determinants of driver injury severity under various control measures.

2.6 Driver Injury Severity by Age and Gender

There have been a considerable number of studies on the development of highway-rail grade crossings' safety studies. However, a study which specifically explores highway vehicle driver injury severity conditioned by age and gender influence, given that a highway-rail grade crossing accident has occurred, has received little attention in previous studies. As a result, studies conducted to study driver's injury severity classified by age and gender for highway accidents are reviewed in this portion of the literature review.

There exist several studies examining significant differences in accident injury severities between different age groups. Abdel-aty (1998) analyzed the effect of driver age on traffic accident on roadway intersections using log-linear models. This model was developed to help understand the relationship between driver age and several important factors including injury severity, average annual daily traffic, roadway character, speed ratio, alcohol involvement, and accident location using an accident database with accidents between 1994 and 1995 in Florida. Findings show that older and very old drivers are more likely to be fatality in traffic accidents due to the decline in their physical condition. Furthermore, very old drivers have a tendency of being involved in angle and turning accidents due to their slower perception and reaction times, and declined ability to judge the speed of oncoming vehicles.

Dissanayake (2002) analyzed factors influential affecting the injury severity of older drivers in passenger car crashes using binary logistic regression models. The sources of data were police crash reports from the state of Florida. Travel speed, use of alcohol and drugs, personal condition, gender, urban/rural nature and grade/ curve existence of the crash location were found to be the important factors impacting the injury severity of older drivers involved car crashes. Higher speeds increase the possibility of an increase in severity for older drivers. If they are not in good physical condition, there is a high likelihood of having an increase in severity for older drivers. Older male drivers when involved in crashes have a higher probability of a lower severity compared to female drivers. Rural locations with curves or grades have a higher probability of generating an increase in severity.

Islam (2006) studied the differences in injury severity between male and female drivers across the different age groups in single-vehicle accidents. The age of the vehicle was also included as a study variable. Separate male and female multinomial logit models describing injury severity were estimated for the young (16 to 24 years), middle-aged (25 to 64 years) and older age vehicle drivers (ages older 65). Findings show statistically significant differences between male and female injury severities among different driver ages and age of vehicle. The finding includes the increased likelihood of fatality for young and older male drivers when driving vehicles less than 5 years ; the increased likelihood of injury for middle-aged female drivers while driving vehicles older than 6 years; and the increase in fatality for older males' beyond 65 years. For behavioral differences, young males have higher fatality probabilities when driving with passengers. For middle-aged females, they have higher injury probabilities when they drive vehicles 6 years old and older.

Boufous (2008) analyzed the injury severity for older drivers as a function of environmental, vehicle and driver characteristics. The study used crash data from New South Wales and Australia. A multiple linear regression analysis showed that road type, the presence of complex intersections, road speed limit, driver's error and use of seat belt were significant predictors of injury severity in older people as a result of a traffic crash. Environmental modifications might contribute to a decrease in the severity of injury as a result of road crashes. For instance, the installation of traffic control devices would decrease the severity of injury. In addition, other improvements would improve the safety of older drivers including increased sign luminance, increased reflectivity of road markings, larger sign symbols and better positioning of traffic signs.

Several studies have found significant differences in highway driver's injury severity between males and females. Ulfarsson (2003) studied male and female injury severities in sport-utility vehicle, minivan, pickup and passenger car accidents at highway locations. Separate multinomial logit models of injury severity are estimated for male and female drivers. Injury severity is classified into no injury, possible injury, evident injury, and fatal injury. The estimation results show that there are significant differences between males and females with regard to factors affecting injury severity. Differences in the driver-injury severity magnitude of effects between the male and female drivers were found. An obvious example is that male drivers striking a barrier or guardrail experienced an increase in the probability of no injury severity while female drivers experienced an increase of fatality. The observed male/female differences suggest a combination of behavioral and physiological factors significantly influence driver's injury severity.

Obeng (2011) studied gender differences in injury severity risks at signalized intersections. The study estimates gender models for injury severity risks and finds that driver condition, type of crash, type of vehicle, and vehicle safety features have different effects on females' and males' injury severity. Monthly crash data at signalized intersections in Greensboro, North Carolina from 1999 to 2002 were used in the model. The data file included 7581 crash records at 301 signalized intersections with 17,116 individual drivers or passengers involved. The evidence shows major gender differences with driver condition, seatbelt use and airbag deployment impacting injury severity risks.

CHAPTER 3

METHODOLOGY

Conventional highway-rail grade crossings studies consider accident prediction models to estimate the number of accidents occurring at crossings. However, few research studies estimate the number of crashes by injury severity. Due to the fact that severity level at a highway-rail grade crossing is naturally ordered, an ordered probit model would be suggested in this study. The objective of this chapter of this dissertation proposal is to state the methodology that will be used in achieving the objectives of this dissertation. A model selection study is given in Section 3.1. A brief introduction of potential crash-injury severity models will be presented in Section 3.2. Section 3.3 introduces the ordered probit model to explore the factors which influence driver's injury severity. Section 3.4 provides a procedure to build-up the final model flow chart.

3.1 Model Selection

A driver injury severity prediction model at highway-rail grade crossings was developed to establish the relationship between the injury severity and contributing factors. Since the dependent variable of the model, driver injury severity, is discrete, discrete choice models are chosen as the suitable approach. Three candidate discrete choice models were selected including: a MNL (Multinomial logit) model, a NL (Nested logit) model, and an OP (Ordered Probit) model. The MNL model is selected because it is by far the most widely used discrete choice model. A distinct limitation is a property known as the "Independence of Irrelevant Alternatives (IIA)". The MNL model does not consider the ordinal property of the ordering characteristic for driver injury severity. Jones et al. (2007) discussed two severe problems of MNL models including the IIA property and the independent and identically distribution (IID) assumption. The IIA property neglects heterogeneity which leads to an inferior model specification and a spurious interpretation of the model. The IID is highly restrictive of parameter estimates causing more variable probability estimates to be independent of another variable's involvement.

Moore (2009) stated that the Nested Logit (NL) model could not prevent the possible correlation within "nested" data sets and the involvement of researcher judgment in the nested structure. The IID problem still exists and the NL model does not recognize the influence from different data sets' heterogeneity affecting the parameter estimation. Zhang (2010) studied the advantage of using ordered probit model. The ordered probit model solves the problem of IIA and ordered discrete data property. As a result, the ordered probit model is selected in this study.

3.2 Ordered Probit Model

3.2.1 Ordered Probit Model Formula

The ordered probit model, which models relationships among ranked outcomes, was used to estimate the injury severity in this research. The multinomial logit model was not selected as this model ignores the ordering of the dependent variable. In this study, driver injury severity is the ordered response.

The general specification of the ordered probit model in this study is given by Equation (3.1) (Zhang, 2011):

$$y_i^* = X_i^T \beta + \varepsilon_i \tag{3.1}$$

Where, X_i is a (K*1) vector of observed non-random explanatory variables measuring the attributes of accident victim i, β is a (K*1) vector of unknown parameters and ε_i is a random error term with zero mean and unit variance for the ordered probit model. In addition, the error terms for different outcomes are assumed to be uncorrelated.

The dependent variable in this study, Y is coded as 1, 2,…, J, defined in equation (3.2):

$$Y = \begin{cases} 1 \text{ if } -\infty \le y_i^* < \tau_1 \\ j \text{ if } \tau_{j-1} \le y_i^* < \tau_j \\ J \text{ If } \tau_{J-1} \le y_i^* < \infty \end{cases} \tag{3.2}$$

Where J is the number of driver injury levels, and ι_j is the threshold value to be estimated for each level. The ordered probit model in equation (3) provides the thresholds which would indicate the levels of inclination causing driver injury severity. In addition, the probabilities of Y taking on each of values j=1,…J are equal to:

$$P(Y=1) = \Phi\left(\tau_1 - X_i^T\beta\right)$$

$$P(Y=j) = \Phi\left(\tau_j - X_i^T\beta\right) - \Phi\left(\tau_{j-1} - X_i^T\beta\right) \tag{3.3}$$

$$P(Y=J) = \Phi\left(\tau_{J-1} - X_i^T\beta\right)$$

Where it is the cumulative probability function of a normal distribution. In our case, Y is chosen as the injury severity, which is grouped into three categories including property-damaged only, injury, and fatality.

3.2.2 Ordered Probit Model Estimation

The parameters of the ordered probit models are estimated using a maximum likelihood estimation method which involves the systematic evaluation of the function at different points to find the point at which the function could be maximized. The log likelihood function in equation (4) is the sum of the individual log probabilities

$$L = \sum_{i=1}^{n}\sum_{j=1}^{3}\log(\Phi\left(\tau_j - X_i^T\beta\right) - \Phi\left(\tau_{j-1} - X_i^T\beta\right)) \tag{3.4}$$

3.2.3 Ordered Probit Model Marginal Effects

Marginal effects are estimated in ordered probit models to get the impacts of variables on probability of each injury severity level (Zhang, 2011). For continuous variables, the marginal effect of a variable for injury severity i could be determined by equation (3.5):

$$P\left(Y = i\right) / \partial X = \left[\phi\left(\mu_{i-1} - \beta X\right) - \phi\left(\mu_i - \beta X\right)\right]\beta \tag{3.5}$$

Where it is the standard normal density

For binary variables, the marginal effect of a variable for injury severity i could be determined by comparing the outcome when the variable takes one value with that when the variable takes zero value, while all other variables remain constant.

$$\Delta\left(Y = i / x_n\right) = \Pr\left(Y = i / x_n = 1\right) - \Pr\left(Y = i / x_n = 0\right) \tag{3.6}$$

3.3 Modeling Procedure

This section provides a general procedural approach to estimate and analyze the ordered probit model. An initial model with all the explanatory variables was calibrated. Independent variables with (P-value >0.05) will be removed in order to get the final model. The final model will be developed with model estimation and marginal analysis in Figure 3.1.

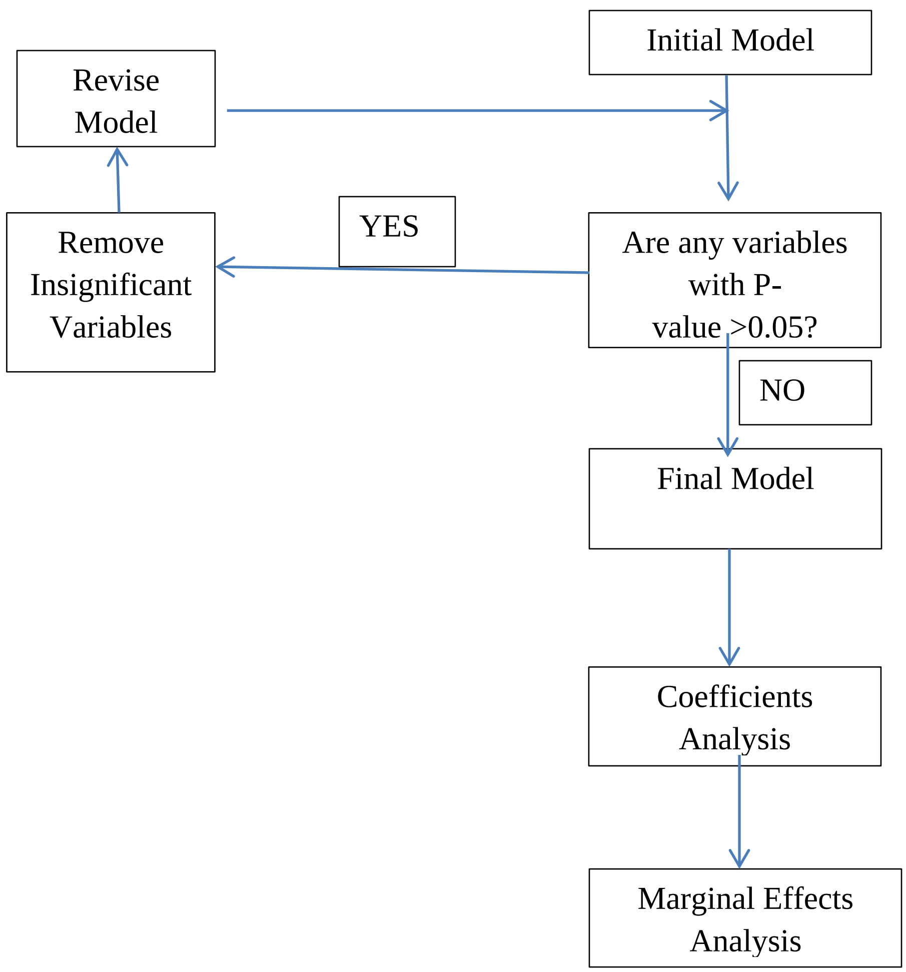


Figure 3.1 Model selection procedure.

CHAPTER 4

DATA PROCESSING

This chapter focuses on the process used for selecting data from available FRA highway-rail grade crossing sources database and the data manipulation procedure to form the sample database. Section 4.1 introduces FRA highway-rail grade crossing database. It will introduce the history of FRA highway-rail grade crossing database and who is responsible for the database. The properties of FRA data will be provided and the classification of FRA database will be given. Section 4.2 details the procedure to how to clean up the data from the FRA database to build our own database. It will detail what types of crashes included in this study and where the data comes from. In addition, it will also provide the detailed data descriptions. Section 4.3 will give the detailed variables correlation matric in order to avoid multicollinearity in our regression study.

4.1 FRA Data Source

The Federal Railroad Administration (FRA) started an original national highway-rail crossing inventory database was on January 1, 1975. The database includes both current and historical records with 80k to 100k crossings updated per year (Woll, 2007). Three sub databases including highway-rail grade crossing inventory, highway-rail crossing history file and highway-rail crossing accident data are classified in the FRA database. The three databases, which are described below, are linked to each other by a common crossing ID number.

Highway-rail grade crossing inventory collects current crossing inventory which reflects the current state of each crossing with reference attributes. It was used to identify independent factors which reflect crossing-related attributes and train/vehicle traffic patterns. In our database, four types of information are obtained: warning device type, area type, AADT, and percentage of trucks. This data are sourced from highway-rail crossing inventory.

The Highway-rail crossing history file reflects the change of the crossings including a reason to update and an effective date of the update. In our study, the highway-rail crossing history file was not utilized.

Highway-rail crossing accident history data provides a history file of accidents which have happened at the crossings and the correlated surrounding conditions at that time. Six types of factors in our final sample database are sourced from highway-rail crossing accident data file including time factors (month, hour, and AM&PM), vehicle information (vehicle speed and vehicle type), train information (train speed), weather information (visibility and weather condition), and driver's information (age, gender, and driver's injury levels).

The data was substantially cleaned and checked for consistency. (i.e. some crossing IDs are missing in the highway-rail crossing inventory but could be found in highway-rail crossing accident data. In this situation, the crossing would not be chosen to be included in the research sample. The overall process of creating the sample database to be used for model estimations comprises the following two steps: (1) highway-rail grade crossing data is extracted from FRA database and (2) Key variable is reclassified in this research. In the first step, the two databases are linked together through the common ID number. The following provides an example of the second step for the variable warning device class at highway-rail grade crossing. This variable contains 9 types of control including no signs or signals, other signs or signals, cross bucks, stop signs, special active warning device, highway traffic signals, flashing lights, all other gates, and four quad gates. The variable is reclassified into three levels: passive control crossings; active control crossings; and no signal control crossings. This classification differs from the highway crossing category because control devices are often implemented together at highway-rail grade crossings (i.e. gates and flashing lights are implemented together as the active control devices).

4.2 Data Formulation

A careful and detailed data collection is essential to obtain reliable conclusions. The original dataset includes 25,945 highway-rail grade crossing accidents from 2002-2011. Finally, 15,881 highway-rail grade crossing accidents were selected as our final research sample after the dataset was cleaned and checked for consistency.

4.2.1 Overall Model Data Formulation

Injury severity is the dependent variable which is ranked as 0-property damaged only, 1-injury, and 2-fatal. The overall model contains 11 variables as shown in Table 4.1. The definition of the variables is also recoded in Table 4.1. The explanatory variables are classified into five groups including “Time factor”, “Weather condition”, “Vehicle and Train Information”, “Environment ”, and “Driver’s Information” in Table 4.1.

Table 4.1 Description of Highway-rail Incidents Characteristics for Analysis

	Variables	Description	Frequency	Percent
Dependent Variable	Injury	0 (Property Damaged only)	10392	65%
		1 (injured)	4037	25%
		2 (fatality)	1419	9%
Time Factor	Peak hour	0 (non-peak)	11127	70%
		1 (peak)	4721	30%
Weather Condition	Unclear weather	0 (clear)	10914	69%
		1 (unclear)	4934	31%
	Dark	0 (other)	11285	71%
		1 (dark)	4563	29%
Vehicle & Train Information	Vehicle speed	0 (Less than 50mph)	15579	98%
		1 (more than 50mph)	269	2%
	AADT	0 (Less 10,000)	13775	87%
		1 (more than 10,000)	2073	13%
	Train speed	0 (less than 50mph)	14270	90%
		1 (more than 50 mph)	1578	10%
Environmental Factors	open space	0 (other areas)	11002	69%
		1 (open space)	4846	31%
	Roadway Pavement	0 (no-paved)	2286	14%
		1 (paved)	13562	85%
Driver's Information	Age	0 (young drivers)	11494	72%
		1 (older than 50 years)	4354	27%
	Gender	1 (Male)	11735	74%
		2 (Female)	4113	26%

4.2.2 Control Device Model Data Formulation

Injury severity is the dependent variable which is ranked as 0-property damaged only, 1-injury, and 2-fatal. For the passive control dataset, the percentage of crashes by the three injury levels is as follows: 63.2% property damaged only, 28.5% injured, and 8.3% fatality. For active control, the percentage of crashes by the three injury levels is as follows: 66.6% property damaged only, 24.1% injured, and 9.3% fatality.

Table 4.2 Description of Control Device Model

	Description	Active Control		Passive Control	
		Number	%	Number	%
Dependent Variable					
Driver	0= property damaged only	6738	66.20%	2674	62.40%
	1= injured	2480	24.40%	1249	29.10%
	2= fatality	961	9.40%	364	8.50%
Independent Variable					
Peak Hour	0 (non-peak)	7338	68.90%	2976	69.40%
	1 (peak)	3316	31.10%	1311	30.60%
Vehicle Speed	0 (more than 50mph)	300	2.90%	118	2.80%
	1 (Less than 50mph)	9879	97.10%	4169	97.20%
Vehicle Type	0 (Other)	8067	79.30%	3142	73.30%
	1 (Truck Related)	2112	20.70%	1145	26.70%
Visibility	0 (dark)	3373	33.10%	858	20.00%
	1 (other)	6806	66.90%	3429	80.00%
Weather	0 (unclear)	3161	31.10%	1573	31.00%
	1 (clear)	7018	68.90%	3506	69.00%
Train Speed	0 (more than 50mph)	1469	14.40%	538	12.50%
	1 (Less than 50mph)	8710	85.60%	3749	87.50%
Driver's Age	0(older than 50 years)	2973	29.20%	1325	30.90%
	1 (young drivers)	7206	70.80%	2962	69.10%
Gender	0 (Male)	7406	72.80%	3317	77.40%
	1 (Female)	2773	27.20%	970	22.60%
Area Type	0 (open space)	2470	24.30%	1998	46.60%
	1 (other areas)	7709	75.70%	2289	53.40%
Roadway Pavement	0 (no-paved)	549	5.40%	1464	34.10%
	1 (paved)	9630	94.60%	2823	65.90%

4.2.3 Age and Gender Data Formulation

Injury severity is the dependent variable which is ranked as 0-property damaged only, 1-injury, and 2-fatal. For the young male dataset, the percentage of crashes by the three injury levels is as follows: 67.2% property damaged only, 25.3% injured, and 7.6% fatality. For middle age male drivers, the percentage of crashes by the three injury levels is as follows: 69.9% property damaged only, 23% injured, and 7.1% fatality.

For old age male drivers, the percentage of crashes by the three injury levels is as follows: 59.3% property damaged only, 24.5% injured, and 16.2% fatality. For the young female dataset, the percentage of crashes by the three injury levels is as follows: 62.6% property damaged only, 30.9% injured, and 6.5% fatality. Estimation of these six unrestricted models is preferable to conducting one restricted model since such individual model allows us to individually investigate the effects of the explanatory variables on injury severity levels by varied age and gender groups. Theoretically the impacts of environmental factors, weather condition, and vehicle and train information on motor vehicle drivers' injury severity are expected to vary across the age and gender groups.

Table 4.3 and Table 4.4 show the frequency and percentage distribution of these variables. For middle age female drivers, the percentage of crashes by the three injury levels is as follows: 62% property damaged only, 30.1% injured, and 7.8% fatality. For old age female drivers, the percentage of crashes by the three injury levels is as follows: 57.2% property damaged only, 29.2% injured, and 13.6% fatality. In addition, the independent variables in this study are made up of continuous variables and categorical variables. Four variables including vehicle speed, train speed, number of lanes, and percent of truck are treated as continuous variables. The remaining five variables include weather, visibility; area type, pavement, and light condition are considered as categorical variables.

Table 4.3 Description of Age and Gender Model for Male

		Young Male		**Middle Male**		**Old Male**	
Dependent Variable							
Injury Severity	0 (Property damaged only)	1800	67.20%	4689	69.90%	1378	59.30%
	1 (Fatality)	203	7.60%	473	7.10%	377	16.20%
	2 (Injured)	677	25.30%	1544	23.00%	569	24.50%
Categorical Variables							
Visibility	0 (Dark)	1173	43.80%	2297	34.30%	530	22.80%
	1 (No-dark)	1507	56.20%	4409	65.70%	1794	77.20%
Weather	0 (No clear weather)	886	33.10%	2095	31.20%	668	28.70%
	1 (Clear weather)	1794	66.90%	4611	68.80%	1656	71.30%
Lights	0 (No)	1562	58.30%	3900	58.20%	1407	60.50%
	1 (Yes)	687	25.60%	1728	25.80%	559	24.10%
Land Use	0 (Open space)	856	31.90%	2173	32.40%	729	31.40%
	1 (None-open)	1823	68.00%	4533	67.60%	1595	68.60%
Pavement	0 (No)	411	15.30%	1051	15.70%	386	16.60%
	1 (Yes)	2269	84.70%	5655	84.30%	1938	83.40%
Continuous Variables	Vehicle Speed	13		10		9	
	Train Speed	30		29		30	
	Traffic lanes	2		2.3		2.3	
	Percent of truck	8		9		9	

There is also description of age and gender for female shown in Table 4.4. The percentage of crashes by the three injury levels is as follows: 62.6% property damaged only, 30.9% injured, and 6.5% fatality. The remaining variables include weather, visibility; area type, pavement, and light condition are considered as categorical variables.

Table 4.4 Description of Age and Gender for Female

		Young Female		Middle Female		Old Female	
		Dependent Variable					
Injury Severity	0 (Property damaged only)	708	62.60%	1305	62.00%	500	57.20%
	1 (Fatality)	74	6.50%	165	7.80%	119	13.60%
	2 (Injured)	349	30.90%	634	30.10%	255	29.20%
		Categorical Variables					
Visibility	0 (Dark)	664	58.70%	1316	62.50%	667	76.30%
	1 (No-dark)	467	41.30%	788	37.50%	207	23.70%
Weather	0 (No clear weather)	781	69.10%	1405	66.80%	645	73.80%
	1 (Clear weather)	350	30.90%	699	33.20%	229	26.20%
Lights	0 (No)	644	56.90%	1218	57.90%	456	52.20%
	1 (Yes)	300	26.50%	530	25.20%	253	28.90%
Land Use	0 (Open space)	323	28.60%	549	26.10%	207	23.70%
	1 (None-open)	808	71.50%	1555	73.90%	667	76.40%
Pavement	0 (No)	131	11.60%	216	10.30%	82	9.40%
	1 (Yes)	1000	88.40%	1888	89.70%	792	90.60%
Continuous Variables	Vehicle Speed	11		9		8	
	Train Speed	28		28		28	
	Traffic lanes	2.3		2.3		2.3	
	Percent of truck	8		8		8	

4.3 Correlation Matrix Studies

In this exercise, correlation matrices are developed using SPSS 16.0 software. In order to avoid multicollinearity in the regression study, the correlation among all independent variables is investigated. Pearson correlation coefficients were estimated to measure the strength of correlation. The Pearson correlation coefficient is usually denoted as *r* and is a value between +1 and -1. The lowest value that *r* can be is 0, this would show zero correlation or no relationship between the two given variables. The highest value that *r* can have is 1.00, this would show a perfect correlation or strong relationship between the two given variables. The values can either be positive or negative. A positive value indicates that an increase in one variable corresponds to an increase in the other variable. A negative value indicates that an increase in one variable corresponds to a decrease in the other variable.

4.3.1 Overall Model Data Correlation Matrix

Injury severity is the dependent variable which is ranked as 0-property damaged only, 1-injury, and 2-fatal. Table 4.5 shows the following sets of predictor variables are not correlated for overall model estimation: schedule factor, vehicle speed (Vsd) , visibility (Vis) , weather condition (Wea), train speed (Tsd), age, gender (Gen), area type, and roadway pavement (Pave).

Table 4.5 Correlation Matrix for Overall Model

	Hour	Vsd	Vis	Wea	Tsd	Age	Gen	Area	Pave
Hour	1	-0.016	-0.185	-0.01	0.029	0.029	0.031	0.018	-0.021
Vehicle speed		1	0.02	0	-0.015	-0.028	-0.019	0.113	0.014
Visibility			1	0.094	-0.026	-0.126	0.013	-0.084	0.113
Weather				1	-0.019	-0.038	0	0.02	0.015
Train speed					1	0.008	-0.027	0.136	-0.128
Age						1	0.001	-0.009	-0.012
Gender							1	-0.055	0.066
Area								1	-0.366
Pave									1

4.3.2 Control Device Model Data Correlation Matrix

Tables (4.5) and (4.6) show the following sets of predictor variables are also not correlated for control device model estimation.

Table 4.6 Correlation Matrix for Active Control Device Model

	Hour	Vsd	Vis	Wea	Train Speed	Age	Gen	Area Type	Pave
Peak Hour	1	0.003	0.215	-0.008	-0.045	-0.04	0.049	-0.01	0.022
Vehicle Speed		1	-0.007	0.017	-0.115	-0.088	0.045	-0.005	0.034
Visibility			1	-0.102	-0.069	-0.18	0.003	0.021	0.055
Weather				1	0.024	0.051	-0.003	0.006	-0.016
Train Speed					1	0.006	0.015	-0.041	-0.136
Age						1	-0.012	0.015	0.008
Gender							1	0.003	-0.034
Area								1	0.077
Pave									1

Table 4.7 Correlation Matrix for Passive Control Device Model

	Peak Hour	Vsd	Vis	Wea	Train Speed	Age	Gen	Area Type	Pave
Peak Hour	1	-0.019	0.115	-0.008	-0.057	-0.011	-0.005	0.005	0.016
Vehicle Speed		1	0.059	-0.029	-0.141	-0.089	-0.008	-0.009	0.089
Visibility			1	-0.08	-0.215	-0.093	-0.034	0.017	0.099
Weather				1	0.041	0.025	0.005	-0.011	-0.018
Train Speed					1	0.006	0.07	-0.129	-0.393
Driver's Age						1	0.041	0.008	-0.002
Gender							1	-0.016	-0.085
Area Type								1	0.158
Pave									1

4.3.3 Age and Gender Model Data Correlation Matrix

From Table 4.8 to Table 4.13, the correlation values show the following sets of predictor variables are also not correlated for driver injury severity model estimation classified by driver's age and gender: vehicle speed, visibility, weather condition, train speed, age, area type, and roadway pavement.

Table 4.8 Correlation Matrix for Young Male Driver's Model

	Vehicle Speed	Visibility	Weather	Train Speed	Area Type	Roadway Pavement
Vehicle Speed	1	-0.035	-0.018	-0.141	-0.01	0.05
Visibility		1	0.016	-0.101	0.039	0.141
Weather			1	0.027	-0.014	0.002
Train Speed				1	-0.036	-0.206
Area Type					1	0.137
Roadway Pavement						1

Table 4.9 Correlation Matrix for Young Female Driver's Model

	Vehicle Speed	Visibility	Weather	Train Speed	Area Type	Roadway Pavement
Vehicle Speed	1	-0.101	0.005	-0.116	0.039	-0.019
Visibility		1	-0.025	-0.064	0.011	0.037
Weather			1	0.022	0.05	-0.009
Train Speed				1	-0.097	-0.126
Area Type					1	0.128
Roadway Pavement						1

Table 4.10 Correlation Matrix for Middle Male Driver's Model

	Vehicle Speed	Visibility	Weather	Train Speed	Area Type	Roadway Pavement
Vehicle Speed	1	0.039	-0.005	-0.142	-0.02	0.044
Visibility		1	0.014	-0.127	0.05	0.122
Weather			1	0.009	-0.014	-0.019
Train Speed				1	-0.076	-0.228
Area Type					1	0.164
Roadway Pavement						1

Table 4.11 Correlation Matrix for Middle Female Driver's Model

	Vehicle Speed	Visibility	Weather	Train Speed	Area Type	Roadway Pavement
Vehicle Speed	1	-0.094	-0.015	-0.102	-0.022	0.019
Visibility		1	0.003	-0.036	0.041	0.109
Weather			1	-0.049	-0.021	-0.034
Train Speed				1	-0.073	-0.187
Area Type					1	0.149
Roadway Pavement						1

Table 4.12 Correlation Matrix for Old Male Driver's Model

	Vehicle Speed	Visibility	Weather	Train Speed	Area Type	Roadway Pavement
Vehicle Speed	1	0.031	-0.004	-0.095	-0.028	0.051
Visibility		1	0.002	-0.139	-0.007	0.104
Weather			1	0.006	0.013	-0.032
Train Speed				1	-0.087	-0.269
Area Type					1	0.176
Roadway Pavement						1

Table 4.13 Correlation Matrix for Old Female Driver's Model

	Vehicle Speed	Visibility	Weather	Train Speed	Area Type	Roadway Pavement
Vehicle Speed	1	-0.053	0.012	-0.113	-0.004	-0.053
Visibility		1	0.043	-0.056	0.043	0.051
Weather			1	-0.018	0.034	0.001
Train Speed				1	-0.017	-0.129
Area Type					1	0.163
Roadway Pavement						1

The intention of this exercise is to determine which variables are not correlated and then use them to develop driver injury severity models. The selection of the models is based on the criteria that variables in the model are not correlated. From the correlation matrices tables from Table 4.5 to Table 4.13, the models satisfy the criterions are:

1. Overall Model: schedule factor, vehicle speed, visibility, weather condition, train speed, age, gender, area type, roadway pavement, and vehicle type.

2. Control Device Model: schedule factor, vehicle speed, visibility, weather condition, train speed, age, gender, area type, and roadway pavement.

3. Age and Gender Model: vehicle speed, visibility, weather condition, train speed, age, area type, and roadway pavement.

CHAPTER 5

MODEL RESULTS AND ANALYSIS

In this section, model results and analysis will be given: overall model results, control device model results, and age and gender model results.

5.1 Overall Model Results

The model was fit using Limdep 9.0 economic software package. The results and model fit information are shown in Table 5.1. The log likelihood value at convergence of the final model is (-1616) and it is significant with a P- value of 0.000.

Table 5.1 Ordered Probit Model Estimation Results

	Estimated Coefficients	**Sig.**
Schedule Factor	-0.111	0.003
Visibility	0.308	0
Weather	0.132	0
Vehicle Type	0.575	0
Vehicle Speed	1.154	0
Train Speed	1.001	0
Area Type	-0.278	0
Pavement	0.314	0
Driver's Age	0.316	0
Gender	-0.18	0
Number of Observations=15,880		
Log likelihood =-1616		
Pseudo R-Square=0.044		
Sig.=0.000		

5.1.1 Model Fit and Estimation Information

A 95 percent confidence interval is used in this study to identify significant variables impacting driver's injury severity at highway-rail grade crossings. The coefficients for the final models are presented in Table 5.1. Coefficients for several sets of explanatory variables in the model are estimated, including "Time factor", "Weather condition", "Vehicle and Train Information", "Environment ", and "Driver's Information"

In this research, schedule factor, or the time the crash occurred, is categorized into two levels: Peak hour and Off-Peak, with peak hour as the reference category. The schedule factor influence is considered given a crash accident has already occurred. From the model results, the coefficient for off-peak is a negative coefficient at -0.111. The negative coefficient indicates that there is a decreased likelihood of higher severities at highway-rail crossings during an off-peak time when compared to accidents happening during the peak hour.

Weather condition is referred to from two aspects: weather and visibility. In this study, the weather factor is classified into two groups: bad weather (such as cloudy, rain, fog, sleet and snow), and clear weather which is selected as the base category. Bad weather has a positive coefficient value of 0.132 which indicates an increased likelihood of severe accidents during bad weather condition at highway rail-grade crossings compared to clear weather condition. Abdel-Aty et al. (2003) found that bad weather conditions make it difficult for drivers to stop or slow down to make a stop. Second, visibility is classified into "other condition" (such as dawn, day, and dusk) and "dark" which is the base category. The positive coefficient value of 0.308 for other or non-dark conditions means an increased likelihood of higher severities for accidents during the other condition. Zhang et.al (2011) found that good light conditions and good weather condition will decrease the probability of severe injuries. The results of this paper show slight differences to what was found by Zhang. The results show that higher severity injuries occurred at highway-rail grade crossings during bad weather and with better visibility.

Highway users' speed describes the driver's estimated speed when the accident occurred. In this research this speed variable is classified into two levels: highway driver's speed more than 50 mph and speed less than 50 mph which is the reference category. The research found speed more than 50 mph was significant with a positive coefficient of 1.154. The positive coefficient indicates an increased likelihood of higher severities at highway-rail crossing injuries for accidents involving vehicular speeds of more than 50 mph when compared to crossing vehicles with speeds less than 50 mph. Zhang et al. (2011) found that the increase of speed limit on freeway will increase the injury severity of the crash.

Railway information here is represented by train speed which describes the estimated train speed when the highway-rail crossing accident occurred. In this research this speed variable is classified into two levels: train speed more than 50 mph and speed less than 50 mph which is the reference category. The research found that speed "less than 50 mph" was significant with a positive coefficient of 1.001. The positive coefficient indicates an increased likelihood of higher severities of highway-rail crossing injuries if the train speed is "more than 50 mph" when compared to train speed "less than 50 mph". A higher train speed means less reaction time for motor vehicle drivers given a highway-rail accident happened and thus increases the probability of higher injury severities at highway-rail crossings. In addition, McCollister et al. (2007) found that increasing train speed will increase injury level which is intuitive.

Vehicle type is classified into two groups: "Truck and Truck-Trailer", and "Auto and other" (other including van, bus, school bus, motorcycle, pedestrian.). "Truck & truck-tra" is chosen as the base category. This research found "auto& other" is significant with a positive coefficient of 0.575. The positive coefficient value implies an increased likelihood of driver injury severity at highway-rail crossing for "auto &other" vehicle drivers when compared to truck drivers. McCollister et al. (2007)'s study found that trucks are mandatory to stop at a highway-rail grade crossing intersections and truck drivers are used to be trained, professional and experienced drivers.

"Area" in this study includes two types: "open space" and "other areas" where "other areas" refer to industrial, commercial, residential and institutional areas. "Open space" is chosen as the reference category. The research found "other area" to be significant with a negative coefficient of -0.278. The negative coefficient indicates a decreased likelihood of higher severities of highway-rail crossing injuries if an accident happens in an area other than open space when compared to open area. This result may be due to driver's lack of alertness and attention while driving in "open space" which may have low traffic volumes. Shankar et al. (1996) in his study on single-vehicle motorcycle accident found that riders' inattention will increase the likelihood of disabling injury in open space area. Zhang et al. (2011) found that accidents located in residential zones will decrease the probability of severe injuries.

Roadways can be paved with timber, asphalt, concrete, rubber, or metal. The roadway pavement in this study is classified as "unpaved" and "paved" which is the reference category. The research found that “unpaved” is significant with a positive coefficient value of 0.314. The positive coefficient value indicates an increased likelihood of higher severities for highway-rail grade crossing accidents if the roadway surface is not paved when compared to a roadway with a paved surface. This could be attributable to the friction level of the roadway. An unpaved road has a lower friction force and therefore needs much more time to stop. As a result, an unpaved roadway will increase the probability of higher severities at highway-rail crossings.

Among the driver’s information, age has a significant effect on injury severities. However, the relationship between driver’s age and injury severity differs by age group. Age in this study is classified into two categories: “less than 50” and “over 50”. This category is based on Abdel-Aty et al. (2003) and Zhang et al. (2011) who looked at injury severity for highway vehicle accidents. “Age less than 50” is defined as the reference category. The research found “over 50” to be significant with positive coefficient 0.316. The positive coefficient value implies an increased likelihood of higher severities for highway-rail crossing injuries for accidents involving older drivers. Furthermore, although older drivers may tend to drive at lower speeds and less likely to be in an accident, once in an accident they tend to have severe injuries by Shankar et al. (1996) and Pai et al. (2007).

Gender is an important factor influencing driver’s injury severity. Female is defined as the reference category. The study found that the variable “male” is significant with a negative coefficient -0.18. The negative coefficient value implies a decreased likelihood of higher severities for highway-rail crossing injuries for accidents involving male drivers when compared to female drivers. Due to physiological differences, women are expected to sustain more severe injuries than men by Yan et al. (2011) and Kockelman et al. (2001).

5.1.2 Overall Model Marginal Effects Analysis

The coefficients estimation in the previous section do not directly reflect the impact of contributing factors on each of the three types of injury levels: property damage only (PDO), injured, and killed. As a result, a marginal effects analysis of factors was conducted. The results in Table 5.2 illustrate the impact of contributing factors on each injury severity level. The coefficient values are classified as positive and negative. A positive marginal coefficient of a variable for a particular injury severity level means that the probability of the severity level will increase as the input variable increases by one unit. The marginal effects of ordered probit model in our study are determined using Limdep 9.0.

Table 5.2 Ordered Probit Model Marginal Effects Analysis Results

	Property Damage Only	**Injured**	**Killed**
Schedule Factor	-0.0088	0.0053	0.0035
Visibility	0.0793	-0.0489	-0.0304
Weather	0.0383	-0.0232	-0.015
Driver's Age	-0.0579	0.0339	0.024
Gender	-0.0002	0.0001	0.0001
Area Type	-0.0003	0.0001	0.0002
Pavement	0.1594	-0.0843	-0.075
Vehicle Type	0.138	-0.0877	-0.0503
Vehicle Speed	-0.273	0.1163	0.1566
Train Speed	-0.2266	0.1114	0.1152

From Table 5.2, the accident occurred during the “Peak hour” will increase the probability of a driver being fatality by 0. 35% and a driver injury by 0.53% compared with “no-peak”. Pai et al. (2007) found that the risk of a severe injury and fatality is higher during the peak period compared with off peak period for motor vehicle drivers in highway collisions.

The variable “bad weather” condition includes cloudy, rain, fog, sleet, and snow. The bad weather condition will increase the probability of “property damage only accidents” by 3.83% compared with clear day condition; on the contrary, it will decrease the probability of injured level accidents by 2.32% and fatality level accidents by 1.5%. This could be explained by the fact that highway vehicle drivers may travel at lower speeds under bad weather condition. This is consistent with the results stated by Duncan et al. (1998) who stated that injury severity was significantly lower on icy or snowy road condition due to slower speeds, maintaining longer headways, and using more caution. The visibility level “dark” was found to decrease the probability of a driver being fatality by 3.04% and decrease the probability of the driver being injured by 4.89% compared with clear condition, whereas it will increase “property damage only level” accidents by 7.93%.

Drivers older than 50 years are more likely to be “injured” or “fatality” in a highway-rail grade crossing accident when compared to drivers that are younger than 50 years. From Table 5.2, drivers older than 50 years will increase the probability of being injured by 3.39 percent and fatality by 2.4 percent compared with drivers younger than 50 years. The increase of the probability of being injured and fatality can be explained by studies which have shown that crash severity increases with age. Abdel-Aty et al. (2003) found that older drivers have a higher probability of more severe injuries especially for drivers above 80 years old. In addition, male drivers will decrease the probability of being fatality and injured. Abdel-Aty et al. (2003) also indicated that female drivers have a higher probability of higher severities.

An accident occurring in an "Open space" area will increase the probability of the driver being fatality by 0.02% and the driver being injured by 0.01% compared with an accident occurring in residential, commercial, and industrial areas. Similarly, Abdel-Aty et al. (2003) found that rural area had a positive influence to increase the probability of driver injury severities. In addition, an accident occurring on at a crossing with "Paved road" will decrease the probability of the driver being fatality by 7.5% and the driver being injured by 8.43% compared with "unpaved" road.

Highway vehicle drivers' with a crossing speed of more than 50 mph is found to increase the probability of a driver being fatality by 15.66% and the driver being injured by 11.63% compared with vehicle drivers with speeds less than 50 mph. ". Abdel-Aty et al. (2003) found that speed increased the probability of severe injuries. For vehicle information," Auto and other" will increase the probability of driver fatality level accidents by 5.03% and driver injured level accidents by 8.77% compared with truck related drivers. This can be explained by the fact that truck drivers are professional and experienced drivers by McCollister et al. (2007). In addition, truck drivers are required to stop at a highway-rail grade crossing regardless of the state of the crossing device.

Train speeds greater than 50 mph was found to increase the probability of a driver being fatality at highway-rail grade crossing accidents by 11.52% and injured by 11.14% compared with a lower train speed. Drivers need to have minimal reaction time to stop once an oncoming train is detected. If the train is coming too fast to cross the highway-rail crossing, highway vehicle drivers will not have enough time to stop and it will significantly increase the likelihood of “fatality level” accidents and “injured level” accidents. McCollister et al. (2007) found that increasing train speed had more effect on injuries and even greater effect on fatalities given that a highway-rail grade crossing accident occurred.

5.2 Control Device Model Results

Ordered Probit models are proposed to be used to analyze the driver injury severities under various control devices for accidents at highway-rail grade crossings. Two Ordered Probit Models are estimated in this study to estimate driver injury severity under active traffic control and passive traffic control. The model examines the effects of explanatory variables on the dependent variable. The estimation model was fit using Limdep 9.0 economic software package. A positive sign of the estimated parameters implies increased injury severities by highway vehicle drivers with increase in the value of the explanatory variables. The P-value for each variable is also listed next to the independent variables. Significant variables are identified as having a p-value of less than 0.05.

The first model shown in Table 5.3 examines the factors affecting injury severities resulting from a highway-rail grade crossing incident controlled by active control devices such as flashing lights and gates. The log likelihood value for the model is (-855) and the P-Value is 0.0 which indicates a good-fit of the model. Factors found to be most significantly associated with the increased injury levels include: weather condition, visibility, vehicle speed, train speed, vehicle type, driver’s age and gender, pavement, and area type.

The second model explores the determinants of driver injury severity resulting from a highway-rail crossing incident controlled by passive control devices such as crossbucks and stop signs. The log likelihood value for the model is (-751) and the P-value is 0.0 which indicates a good-fit of the model. Model estimation results indicate that schedule factor, visibility, vehicle speed, vehicle type, train speed, driver's age and gender, area type, and pavement are significant variables associated with driver injury severity as shown in Table 5.3. The following section provides a more detailed discussion of these findings.

Table 5.3 Control Device Model Estimation Results

Variables	**Active Control**		**Passive Control**	
	Parameter Estimate	**P-Value**	**Parameter Estimate**	**P-Value**
Peak hour	/	/	-0.169	0.014
Weather	0.123	0.008	/	/
Visibility	0.366	0	0.177	0.039
Vehicle speed	1.215	0	0.966	0
Train speed	1.021	0	0.885	0
Age	0.345	0	0.284	0
Area Type	/	/	0.29	0
Pavement	0.353	0	0.266	0
Number of Observation	10194		5079	
Log likelihood	-855		-751	
Pseudo R-Squared	0.047		0.045	
Significance (P-value)	0		0	

5.2.1 Control Device Model Estimation Results

A 95 percent confidence interval is used in this study to identify significant variables impacting driver's injury severity at highway-rail grade crossings. The coefficients for the final models are presented in Table 5.3.

In this study, schedule factor is categorized into two levels: peak hour and off-peak, with off-peak as the reference category variable. The schedule factor is significant only for passive control highway-rail crossings. There is an increased likelihood of higher severities at highway-rail crossing injuries for accidents happening during the peak hour under passive control when compared to accidents happening during the peak hour under active control.

Weather factor is classified into two groups: bad weather (such as cloudy, rain, fog, sleet and snow) and good weather which is selected as the reference category. Weather is found to be a significant variable influencing highway driver's injury severity only under active control highway-rail grade crossing intersections. The positive coefficient value (0.123) implies that drivers under active control are more likely to have a severe injury in bad weather condition. Visibility is classified into "dark" and "other condition" (such as dawn, day, and dusk) which is the reference category. Drivers are found to have severe injuries under "dark" condition at active control highway-rail grade crossings.

Highway users' speed describes the driver's estimated speed when the accident occurred. In this research this speed variable is classified into two levels: highway driver's speed "more than 50 mph" and speed "less than 50 mph" which is the reference category. The coefficient estimates for the highway driver's speed among the two models indicate a difference by type of control. Both the passive and actively controlled crossings have a positive coefficient for the highway driver's speed which means an increased likelihood of more severe highway-rail crossing injuries with increasing speed. There is an increased likelihood of higher severities at highway-rail crossings with high speed under active control with a coefficient estimate of (1.215) when compared to accidents happening with high speed under passive control with a coefficient estimate of (0.966).

Railway information here is represented by train speed which describes the estimated train speed when the highway-rail crossing accident occurred. In this research this speed variable is classified into two levels: train speed "more than 50 mph" and speed " less than 50 mph" which is the reference category. There is an increased likelihood of higher severities at highway-rail crossings for accidents happening under high train speed and active control with a coefficient estimate of (1.021) when compared to accidents happening under high train speed and passive control with a coefficient estimate of (0.885).

Among the driver's information, age has a significant effect on injury severities. However, the relationship between driver's age and injury severity differs by age group. Age in this study is classified into two categories: "less than 50" and "over 50". This category is based on Abdel-Aty et al. (2003) and Zhang et al. (2011) who looked at injury severity for highway vehicle accidents. "Age less than 50" is defined as the reference category. The coefficient estimate for active control (0.345) is larger than the coefficient value (0.284) under passive control which implies that older drivers are more likely to have severe injury accidents under active control compared to older drivers under passive control given a highway-rail accident has occurred. This could be explained by the fact that older drivers have slower reactions compared to younger drivers. This findings could be supported by Pai (2007)'s study. In his study, Pai found that drivers over 60 years are less likely to have severe driver injuries at stop, give-way signs or markings controlled junctions.

"Area" in this study includes two types: "open space" and "other areas" where "other areas" refer to industrial, commercial, residential and institutional areas. "Open space" is chosen as the reference category. "Area type" is found to be a significant variable to influence highway driver's injury severity only under passive control at highway-rail grade crossing intersections. The coefficient estimate for passive control (0.29) implies that drivers in open space area are more likely to have severe injury accidents given a highway-rail accident has happened than compared to other areas.

Pavement here is classified as "highway unpaved" and "highway paved" with "highway paved" being the reference category. The coefficient estimate under active control of (0.353) is larger than the estimate under passive control with a coefficient of (0.266). This implies that highway drivers on "unpaved" highways under active control devices are more likely to have severe injuries compared with passive control highway-rail crossings given an accident already happened.

5.2.2 Control Device Model Marginal Effects Analysis

Table 5.4 indicates the marginal effects of significant independent variables on the probabilities of each injury severity level (Y=0 uninjured; Y=1 injured; Y=2 killed). In addition, the marginal effects analysis is considered under both active control and passive control. The definition of marginal effects could be described as the increased or decreased probabilities in each injury severity level associated with the change of significant independent variables. For categorical variables, the marginal coefficients reflect the change of probability of injury severity compared to the reference categorical variable when all other independent variables remain the same.

Table 5.4 Control Device Model Marginal Effects Analysis Results

		Injury Level	**Active Control**	**Passive Control**
Time Factor	Peak hour	Y=0	/	0.0083
		Y=1	/	-0.005
		Y=2	/	-0.0032
Weather	Weather	Y=0	-0.043	/
		Y=1	0.0253	/
		Y=2	0.0177	/
	Visibility	Y=0	-0.0608	-0.0942
		Y=1	0.0359	0.06
		Y=2	0.0249	0.0342
Train and Vehicle Information	Vehicle speed	Y=0	-0.111	-0.2389
		Y=1	0.0577	0.1059
		Y=2	0.0533	0.133
	Train speed	Y=0	-0.2051	-0.1787
		Y=1	0.1004	0.0923
		Y=2	0.1047	0.0864
Driver's Information	Age	Y=0	-0.0845	-0.0258
		Y=1	0.0473	0.0155
		Y=2	0.0372	0.0103
Area Type	Area Type	Y=0	/	-0.0159
		Y=1	/	0.0096
		Y=2	/	0.0062
	Pavement	Y=0	-0.0169	-0.1814
		Y=1	0.0099	0.1026
		Y=2	0.0069	0.0788

For example, for the categorical variable pavement, compared to a highway-rail grade crossing accident on a paved highway, a highway-rail grade crossing accident on an unpaved highway will increase the probability of injury accidents by 0.99%, and fatality level accidents by 0.69%, while decreasing the probability of property damaged only level accidents by 1.69% at active control highway-rail grade crossings. At highway-rail grade crossings with passive control, however, a highway-rail grade crossing accident on unpaved highway will increase the probability of the probability of injured level accidents by 10.26%, and fatality level accidents by 7.88% while decreasing the probability of property damaged only level accidents by 18.14% at passive control highway-rail grade crossings. A conclusion could be made that unpaved roads result in greater severities for drivers at passive control than active control, when all other independent variables remain the same.

5.3 Age and Gender Model Results

Ordered Probit models are used to analyze driver injury severities under various drivers' age and gender groups for the highway-rail grade crossing collisions. Age classification is based on Islam's study in the year 2006 and a total of six models are estimated: young male drivers (ages 15 to 24) as model 1, young female drivers (ages 15 to 24) as model 2, middle- aged male drivers (ages 25 to 55) as model 3, middle-aged female drivers (ages 25 to 55) as model 4, older male drivers (ages 56 and older) as model 5, and older female drivers (ages 56 and older) as model 6. Tables (5.5) and (5.6) show the coefficient estimation for the six models. Table 5.5 show model results for male drivers and Table 5.6 show model results for female driver. Table 5.7 shows the marginal effects analysis for the age and gender models. The model estimation results for all models are reported in this section first, followed by individual model discussions. A positive sign of the estimated parameters implies increased injury severities by highway vehicle drivers with increase in the value of the explanatory variables. The P-value for each variable is also listed next to the independent variables in Tables 5.5 and 5.6. Significant variables are identified as having a p-value of less than 0.05.

The young male model (Model 1) examines the factors which impact injury severities for young male drivers (ages 15 to 24). Factors found to significantly impact driver injury include: vehicle speed, train speed, visibility, weather condition, and roadway pavement.

The middle group male model (Model 2) examines the factors which impact injury severities for middle age male drivers (ages 25 to 55). In this model, the factors which are shown to significantly influence driver injury severity include vehicle speed, train speed, visibility, weather condition, and roadway pavement.

The Older male model (Model 3) examines the factors which impact injury severities for older male drivers (age 56 and over 56 years old). A variety of the explanatory variables are found to be statistically significant for older male drivers' injury severity: vehicle speed, weather condition, train speed, area type and roadway pavement.

Young female model (Model 4) explores the factors influencing young female drivers' injury severity (age 15 to 24 years old). The factors found to significantly influence female drivers' injury severity include: vehicle speed, visibility, and train speed.

The middle group female model (Model 5) examines the factors which impact injury severities for middle age female drivers (age 25 to 54 years old).

Table 5.5 Male Model Estimation Results

	Young Aged Male	P-Value	Middle Aged Male	P-Value	Old Aged Male	P-Value
Vehicle speed	0.016	0	0.017	0	0.016	0
Dark Condition	0.233	0	0.095	0.006	/	/
Bad Weather	0.178	0.001	0.14	0	0.126	0.026
train Speed	0.014	0	0.015	0	0.024	0
Open Area	/	/	/	/	-0.513	0
Unpaved Road	0.437	0	0.393	0	0.259	0
Number of Observation	2680		6706		2324	
Log likelihood	-2051		-5199		-2000	
Pseudo R-Squared	0.055		0.06		0.093	
Significance (P-value)	0		0		0	

Table 5.6 Female Model Estimation Results

	Young Aged Female	P-Value	Middle Aged Female	P-Value	Old Aged Female	P-Value
Vehicle speed	0.024	0	0.022	0	0.02	0
Dark Condition	0.316	0	0.151	0.008	/	/
Bad Weather	/	/	0.143	0.015	/	/
train Speed	0.026	0	0.021	0	0.032	0
Open Area	/	/	/	/	-0.42	0
Unpaved Road	/	/	0.231	0.006	/	/
Number of Observation	1131		2104		874	
Log likelihood	-828		-1634		-719	
Pseudo R-Squared	0.123		0.094		0.133	
Significance (P-value)	0		0		0	

Older female model (Model 6) examines the factors which impact injury severities for older female drivers (age 56 and over 56 years old). In this model variables found to significantly impact driver injury severity include: vehicle speed, train speed, and area type. The following section provides a more detailed discussion of these findings.

5.3.1 Age and Gender Model Estimation Analysis

A 95 percent confidence interval is used in this study to identify significant variables impacting driver's injury severity at highway-rail grade crossings. The coefficient estimations for the final models are presented in Table 5.5 and 5.6.

Visibility is categorized into "dark" condition and "non-dark" condition which is the reference category. "Dark" visibility condition has a positive coefficient value for all age groups which means an increased likelihood of severe incidents during "dark" conditions at highway rail-grade crossings compared to "non-dark" condition. As shown in Table 5.5 for male drivers, the research found "dark" condition was significant with positive coefficients for young male drivers (0.233) and middle age male drivers (0.095) but is not significant for older male drivers. These differing coefficient values indicate that young male drivers are more likely to be influenced by "dark" condition compared with middle age male and older male drivers.

Second, for female drivers as shown in Table 5.6, the study found that "dark" condition was significant with positive coefficients for young female drivers (0.316) and middle age female drivers (0.151) but was not found to be significant for older female drivers.

Third, for young drivers across two gender groups (Model 1&4), the study found “dark” condition was significant with positive coefficients for young female drivers with a coefficient of (0.316) which is greater than the coefficient value of (0.178) for young male drivers. This implies that “dark” condition influences the injury severity of young female drivers than young male drivers.

Fourth, for middle age drivers across two gender groups, the study found “dark” condition was significant with positive coefficients for middle age female drivers with a coefficient of (0.151) which is greater than the coefficient value for middle age male drivers of (0.095). This implies that “dark” condition influences the injury severity of middle age female drivers more than middle age male drivers. As a conclusion, the injury severity of young drivers is more likely to be influenced by visibility compared to middle age and older drivers. In addition, the injury severity of female drivers is found to be more influenced by visibility compared with male drivers based on previous discussion.

The weather condition is grouped two categories: bad weather (such as cloudy, rain, fog, sleet and snow), and clear weather which is selected as the reference category. Bad weather has a positive coefficient value for all age groups which means an increased likelihood of severe accidents during bad weather condition at highway rail-grade crossings compared to clear weather condition. As shown in Models 1, 2 and 3 of Table 5.5, for male drivers , the coefficient value for bad weather for young male drivers is positive (0.178) and is greater than the coefficient for middle age drivers (0.14) and older male drivers (0.126). This indicates that injury severities for young male drivers are more likely to be influenced by bad weather condition (such as cloudy, rain, fog, sleet and snow) compared with middle age male drivers and older drivers. Second, Models 4, 5 and 6 in Table 5.6 for female drivers, shows that the variable bad weather is significant only for middle age female drivers. As a result, accidents at highway rail-grade crossings under bad weather condition are more likely to result in more severe driver injuries for male drivers compared to female drivers.

Vehicle speed is the continuous variable describing speed of vehicle in miles per hour. As shown in Tables (5.5) and (5.6), the coefficient estimation of vehicle speed for young female drivers is positive (0.024) which is greater than the coefficient for young male drivers (0.016). Second, the coefficient estimate for middle age female drivers is positive (0.022) which is greater than the coefficient estimate for middle age male drivers (0.017), as shown in models 2 and 5. Third, based on model comparisons between models 3 and 6, the coefficient estimate for older female drivers is positive (0.02) which is greater than the coefficient estimate for older male drivers (0.016). Moreover, the coefficient values for vehicle speed among male drivers by different age groups are almost the same. This is also true for female drivers by different age groups.

Train speed is the continuous variable describing speed of train in miles per hour. For all age groups, drivers are found to have more severe injuries with increasing train speeds. The coefficient estimate for train speed for older male drivers is positive (0.024) and is greater than the coefficient for young male drivers (0.014) and middle age drivers (0.015). This implies that an increase of train speed, given an accident has occurred, will more likely cause a severe injury for older male drivers compared with young male drivers and middle age male drivers. For female drivers, Models 4, 5 and 6, the coefficient estimate for vehicle speed for middle age female drivers is positive (0.021) which is lower than the coefficient for young female drivers (0.026) and older female drivers (0.032). This implies that an increase of train speed, given an accident has occurred, will less likely cause a severe injury for middle age female drivers compared with young female drivers and older female drivers. In addition, the coefficient estimate for young female drivers is positive (0.026) which is greater than the coefficient estimate for young male drivers (0.014). The coefficient estimate for middle age female drivers is positive (0.021) which is greater than the coefficient estimate for middle age male drivers (0.015). The coefficient estimate for older female drivers is positive (0.032) which is greater than the coefficient estimate for older male drivers (0.024). This implies that an increase of train speed, given an accident has occurred, will more likely cause a more severe injury for female drivers compared to male drivers.

Roadway pavement is categorized into two levels: unpaved and paved, with paved as the reference category variable. The coefficient estimates indicate more severe injuries if the accident occurred on a roadway that is unpaved. For male drivers (Model 1, 2 and 3), the coefficient estimate for roadway pavement for older male drivers is positive (0.024) which is greater than the coefficient estimate for young male drivers (0.014) and middle age drivers (0.015). This implies that unpaved roadway, given an accident has occurred, will more likely cause a severe injury for older male drivers compared with young male drivers and middle age male drivers. However, the coefficient values for female drivers are not clear. As a result, an unpaved roadway is more likely to result in higher injury severities for male drivers compared with female drivers.

"Area" here includes two types: "open space" and "other areas" where "other areas" refer to industrial, commercial, residential and institutional areas. "Open space" is chosen as the reference category. The research found "other area" to be significant with negative coefficients for older male drivers (-0.513) and older female drivers (-0.42). The negative coefficient indicates a decreased likelihood of severe highway-rail crossing injuries if an incident happens in an area other than open space when compared to open area. In addition, older drivers are less likely to have higher injury severities compared with young and middle age drivers in open areas.

5.3.2 Age and Gender Model Marginal Effects Estimation Analysis

Table 5.7 indicates the marginal effects of significant independent variables on the probabilities of each injury severity level (Y=0 property damaged only; Y=1 injured; Y=2 fatal). In addition, the marginal effects analysis is considered under both active control and passive control. The definition of marginal effects could be described as the increased or decreased probabilities in each injury severity level associated with the change of significant independent variables. For continuous variables, the marginal coefficients reflect the change of probability of injury severity by one unit increase of the independent variable, keeping other factors at the same value. For categorical variables, the marginal coefficients reflect the change of probability of injury severity compared to the reference categorical variable when all other independent variables remain the same.

A highway-rail grade crossing accident at dark visibility condition will increase the probability of injured level accidents by 5.62%, and fatality level accidents by 2.2% while decreasing the probability of property damaged only level accidents by 7.82% for young male drivers (Model 1). For young female drivers (Model 4), however, a highway-rail grade crossing accident at "dark" visibility condition will increase the probability of injured level accidents by 9.12%, and fatality level accidents by 2.5%, while decreasing the probability of property damaged only level accidents by 11.61%. Third, for middle age male drivers (Model 2), a highway-rail grade crossing accident at "dark" visibility condition will increase the probability of injured level accidents by 2.23%, and fatality level accidents by 1.02%, while decreasing the probability of property damaged only level accidents by 3.25%.

Table 5.7 Model Marginal for Age and Gender Model

	Injury Level	Young Aged Male	Middle Aged Male	Old Aged Male	Young Aged Female	Middle Aged Female	Old Aged Female
Weather	Y=0	-0.0643	-0.0475	-0.0481	/	-0.0529	/
	Y=1	0.0432	0.0328	0.0224	/	0.0385	/
	Y=2	0.021	0.0147	0.0258	/	0.0144	/
Visibility	Y=0	-0.0782	-0.0325	/	-0.1161	-0.056	/
	Y=1	0.0562	0.0223	/	0.0912	0.0407	/
	Y=2	0.022	0.0102	/	0.025	0.0154	/
Vehicle speed	Y=0	-0.0057	-0.0059	-0.0061	-0.0089	-0.0084	-0.0079
	Y=1	0.0038	0.004	0.0028	0.0069	0.0061	0.0046
	Y=2	0.0019	0.0019	0.0033	0.002	0.0024	0.0034
Train speed	Y=0	-0.005	-0.0053	-0.0093	-0.0097	-0.0077	-0.0127
	Y=1	0.0033	0.0036	0.0042	0.0075	0.0055	0.0073
	Y=2	0.0017	0.0017	0.0051	0.0021	0.0022	0.0054
Pavement	Y=0	-0.1671	-0.1441	-0.1016	/	-0.0884	/
	Y=1	0.0991	0.0902	0.0411	/	0.0604	/
	Y=2	0.068	0.0539	0.0605	/	0.028	/
Open Area	Y=0	/	/	0.1982	/	/	0.2556
	Y=1	/	/	-0.0897	/	/	-0.1476
	Y=2	/	/	-0.1085	/	/	-0.108

In addition, for middle age female drivers (Model 5), a highway-rail grade crossing accident at "dark" visibility condition will increase the probability of injured level accidents by 4.07%, and fatality level accidents by 1.54%, while decreasing the probability of property damaged only level accidents by 5.6%. In conclusion, a result could be made that young drivers are more likely to be influenced by visibility compared with middle age and older drivers and female drivers are found to be influenced by visibility compared with male drivers.

Highway driver's speed is a continuous variable. For young male drivers (Model 1), a 10 mph increase in highway driver's speed will increase the probability of injured level accident by 0.38%, and fatality level accident by 0.19%, while it will decrease the property damaged only level accident by 0.57%. For young female drivers (Model 4), however, a 10 mph increase in highway driver's speed will increase the probability of injured level accident by 0.69%, and fatality level accident by 0.2%, while it will decrease the property damaged only level accident by 0.89%. For middle age male drivers (Model 2), a 10 mph increase in highway driver's speed will increase the probability of injured level accident by 0.4%, and fatality level accident by 0.19%, while it will decrease the property damaged only level accident by 0.59%. For middle age female drivers (Model 5), however, a 10 mph increase in highway driver's speed will increase the probability of injured level accident by 0.61%, and fatality level accident by 0.24%, while it will decrease the property damaged only level accident by 0.85%. For older male drivers (Model 3), a 10 mph increase in highway driver's speed will increase the probability of injured level accident by 0.28%, and fatality level accident by 0.33%, while it will decrease the property damaged only level accident by 0.61%. For older female drivers (Model 6), however, a 10 mph increase in highway driver's speed will increase the probability of injured level accident by 0.46%, and fatality level accident by 0.34%, while it will decrease the property damaged only level accident by 0.79%. As a conclusion, it could be said that an increase of vehicle speed, given an accident has occurred, will more likely cause for female drivers compared with male drivers.

CHAPTER 6

CONCLUSIONS AND FUTURE WORKS

An ordered probit model is introduced in this study to analyze the factors influencing driver's injury severity at highway-rail crossings. Three model conclusions are summarized and followed by future work studies.

6.1 Overall Model Conclusion

An ordered probit model was developed in this study to analyze the factors influencing driver's injury severity at highway-rail crossings. The model was developed using accidents from 2002-2011 locations all over the United States. As a result, the research uses a dataset which is the latest and comprehensive data file. Analysis of the ordered probit model in our study reveals crucial factors influencing highway driver's injury severity, and it will also provide potential strategies to reduce driver injury severity at highway-rail grade crossings. Based on the model estimation and marginal analysis results, it was found that the factors significantly impacting driver injury severity include peak hour, weather, visibility, vehicle type, vehicle speed, train speed, driver's age, gender, area type and highway pavement. Marginal analysis was provided to quantitatively explain the marginal effects of each independent variable on each injury level.

The study found that female drivers are more likely to have an increase in severity at highway-rail crossings compared to male drivers. Older drivers are more susceptible than younger drivers to cause an increase in severity at highway-rail crossings. An increase in severity is more likely under bad weather road condition, such as wet, icy or snowy road surface, and by visibility, such as dark conditions. In addition, a reduced speed limit for train and vehicles will significantly reduce driver injury severity.

Although previous researchers have focused on analyzing the frequency of crashes at highway-rail grade-crossings, few studies have been conducted on driver's injury severity level. In addition, previous driver injury level studies at highway-rail grade crossing did not account for the ordered nature of injury levels (Miranda-Moreno, 2009; Hu, 2009; McCollister, 2007). This research attempted to identify contributing factors which influenced the incident driver's injury severity at highway-rail grade crossings. This study provides differences in methodology and dataset resulting in a contribution to this field of safety of highway-rail crossings. The findings are beneficial to transportation engineers to improve safety performance at highway-rail grade crossings.

Further studies should be performed to address the limitations of this study. The assumption of this study suggests that the input variables are independent among each other. Highway driver's information is found to be significant variable to influence driver injury severity at highway-rail grade crossings. As a result, more driver information, such as use of alcohol and educational status, should also be collected to provide more drivers' information.

6.2 Control Device Model Conclusion

Utilizing the most recent ten years (2002-2011) of highway-rail grade crossing accidents, results of two ordered probit models in this study uncovered crucial determinants of highway driver injury severity at highway-rail grade crossings under different control measures. The findings offer insights into potential prevention strategies which could be undertaken to reduce driver injury severities.

Based on the model estimation and marginal analysis results, it was found that the factors significantly impacting driver injury severity at highway-rail grade crossings include peak hour, visibility, vehicle speed, train speed, percent of truck, driver's age, and highway pavement for both active and passively controlled highway-rail grade crossings. A marginal analysis was provided to quantitatively explain the marginal effects of each independent variable on each injury level.

The analysis of driver injury severity under various control devices could help reduce the severity of accidents at highway-rail grade crossings and increase driver's safety. The detailed findings are now listed by active or passive control. For active control highway-rail grade crossings where there are high volumes of trains and vehicles, speed reduction for both trains and vehicles will significantly reduce driver injury severity. In addition, paving highways at highway-rail grade crossings will also help to reduce driver injury severity at highway-rail crossing accidents. Highway driver's age, weather condition and visibility also work as important factors influencing driver injury severity at highway-rail crossings.

For passive control highway-rail grade crossings, vehicle speed and train speed are also found to be crucial to influence highway driver's injury severity. However, the level of influence by vehicle speed and train speed at passive control is lower compared with active control. Pavement, weather condition, and visibility are found to play a much more important role compared to active control. As a recommendation, improving highway pavement will significantly reduce driver injury severity at passive control highway-rail grade crossings. In addition, drivers should pay more attention while crossing passive control highway-rail grade crossings under bad weather conditions.

In summary, this study explored the contributing factors to driver injury severity at both passive control crossings and active control crossings. The findings are beneficial to transportation engineer to address highway-rail grade crossing safety problem at various control devices. However, this study does suffer from several limitations. Further study is needed to investigate combination of factors, such as whether driver's age and gender work together to influence driver's injury severity. In addition, more driver information, such as alcohol use and educational status, should also be collected to provide more drivers' information.

6.3 Age and Gender Conclusion

The purpose of this study is to explore the differences in driver-injury severity between male and female drivers and across three age groups for highway-rail grade crossing accidents involving vehicle drivers. Studying highway-rail grade crossing accidents from 2002-2011, six separate ordered probit models are estimated. Model estimation is conducted to evaluate the differences between different age and gender groups and finally a marginal analysis was performed and the results compared between models.

For male drivers, Vehicle speed, train speed, weather condition, and roadway pavement are four common variables across all the three age groups. However, there are differences existing among the three age groups. First, young male drivers are more likely to be influenced by "dark" condition compared with middle age male drivers and older male drivers. Second, young male drivers are more likely to be influenced by bad weather condition (such as cloudy, rain, fog, sleet and snow) compared with middle age male and older male drivers. Third, an increase of train speed, given an accident has occurred, will more likely cause a more severe injury for older male drivers compared with young male drivers and middle age male drivers. For female drivers, Vehicle speed and train speed are two common variables across all the three age groups. Initially, the visibility coefficient for young female drivers is greater than the coefficient estimate for middle age female drivers which may indicate that drivers are driving more carefully as their age increases under “dark” condition. In addition, an increase of train speed, given an accident has occurred, will less likely cause a severe injury for middle age female drivers compared with young female drivers and older female drivers.

For young age drivers, vehicle speed, visibility, and train speed are three common variables across the two gender groups. However, there are differences existing among the young age gender groups. First, visibility “dark” condition influences young female drivers than young male drivers. Second, an increase of train speed, given an accident has occurred, will more likely cause a severe injury for young female drivers compared with young male drivers. Third, an increase of vehicle speed, given an accident has occurred, will more likely cause a more severe injury for young female drivers compared with young male drivers.

For middle age drivers, vehicle speed, visibility, weather condition, train speed, and roadway pavement are five common variables across the two gender groups. However, there are differences existing among middle age gender groups. First, visibility “dark” condition influences middle age female drivers more than middle age male drivers. Second, an increase of train speed, given an accident has occurred, will more likely cause a severe injury for middle age female drivers compared with middle age male drivers. Third, an increase of vehicle speed, given an accident has occurred, will more likely cause a severe injury for middle age female drivers compared with middle age male drivers.

For older drivers, vehicle speed, train speed and area type are the three common variables across the older gender groups. An increase of vehicle speed, given an accident has occurred, will more likely cause a severe injury for older female drivers compared with older male drivers. In addition, an increase of train speed, given an accident has occurred, will more likely cause a severe injury for older female drivers compared with older male drivers.

In summary, this study explored the contributing factors to driver injury severity between male and female drivers and across three age groups. The findings are beneficial to transportation engineer to address highway-rail grade crossing safety problem for varied type of vehicle drivers. However, this study does suffer from several limitations. Estimation of injury severity models separately analyzes explanatory variables on injury severity by genders across different age groups. Further study is need to investigate comprehensive driver's information, such as driver's age and gender related biomechanics and behavioral attributes including driver's height and weight. In addition, more environmental factors, such as vegetation clearance at the highway-rail grade crossing, should also be collected to provide more drivers' information.

6.4 Summary and Future Studies

The research shows there are differences in the factors which influence motor vehicle driver's injury severity given a highway-rail grade crossing accident happened. These differences should be considered in the development of transportation government policy or operational changes aimed at reducing driver injury severity. The implication of the results obtained in this research is that older drivers are more susceptible than younger drivers to cause an increase in severity at highway-rail crossings. An increase in severity is more likely under bad weather road condition, such as wet, icy or snowy road surface, and by visibility, such as dark conditions. In addition, improving highway pavement will significantly reduce driver injury severity at passive control highway-rail grade crossings. Furthermore, young male drivers are more likely to be influenced by bad weather condition (such as cloudy, rain, fog, sleet and snow) compared with middle age male and older male drivers.

In this section, future studies will be discussed based on the limitation of the study from three aspects: data source limitation, model assumption, and model itself. First, the primary data source used in this study is the FRA database data file which covered three sub databases including highway-rail grade crossing inventory, highway-rail crossing history file and highway-rail crossing accident data. It included five datasets "Schedule factor", "Weather condition", "Vehicle and Train Information", "Environment ", and "Driver's Information". There is no secondary data source available for this study. Therefore, only driver's age and gender are included. Future studies could look into more drivers' factors such as alcohol use and educational status. The impacts of these factors on highway accidents are discussed in some of the literature, but how they impact driver injury severity for highway-rail grade crossing collisions have not yet been studied.

The assumption of this study suggests that the input variables are independent among each other. The potential correlations between each variable are not considered. Further study is needed to investigate factors interactions, such as driver's age and gender work together to influence driver's injury severity. In addition, more driver information, such as alcohol use and educational status, should also be collected to provide more drivers' information.

For the model choice, the ordered probit model addresses the problem of IIA and ordered discrete data and as a result includes in this study. However, the ordered probit model also suffers from the assumption of a normal distribution for all unobserved components of utility. Therefore, a more flexible model, such as an ordered mixed model, is suggested in the future study. The ordered mixed logit model is a highly flexible model which could approximate any random utility model without assumption that the error terms following a normal distribution.

REFERENCES

Abdel-Aty, M. (2003). "Analysis of Driver Injury Severity Levels at Multiple Locations Using Ordered Probit Models." *Journal of Safety Research*, Vol. 34, No. 5, pp. 597-603.

Abdel-Aty, M., C.L. Chen, and J.R. Schott (1998). "An Assessment of The Effect of Driver Age on Traffic Accident Involvement Using Log-linear Models." *Accident Analysis and Prevention*, Vol. 30, No. 6, pp. 851-861.

Abraham, J., T.K. Datta, and S. Datta (1998). "Driver Behavior at Rail-highway Crossings." *Transportation Research Record*, No. 1648, pp. 28-34.

Anandarao, S. and C. Martland (1998). "Level-Crossing Safety on East Japan Railway Company: Application of Probabilistic Risk Assessment Techniques." *Transportation*, Vol. 25, No. 3, pp. 265-286.

Association of American Railroads (2011). "Statement of Edward R. Hamberger President &Chief Executive Officer." http://www.aar.org/KeyIssues/~/media/aar/Testimony/2011/2011-04-Testimony-SAFETEA-LU.ashx. Accessed in 2012.

Austin, R.D. and Carson, J.L. (2002). "An Alternative Accident Prediction Model for Highway-rail Interfaces." *Accident Analysis and Prevention*, Vol. 34, No. 1, pp. 31-42.

Berg, W., K. Knoblauch, and W. Hucke (1982). "Causal Factors in Railroad-highway Grade Crossing Accidents." *Transportation Research Record*, No. 847, pp. 47-54.

Boufous S., C. Finch, A. Hayen, and A. Williamson (2008). "The Impact of Environmental, Vehicle and Driver Characteristics on Injury Severity in Older Drivers Hospitalized As A Result of A Traffic Crash." *Journal of Safety Research*, Vol. 39, No. 1, pp. 65-72.

Bureau of Transportation Statistics (2011). "Transportation Statistics Annual Report." *U.S. Department of Transportation Research and Innovative Technology Administration*, *http://www.bts.gov/publications/national_transportation_statistics/pdf/entire.pdf*. Accessed in 2012.

Cameron, A.C., and F.A.G. Windmeijer (1995). "An R-Squared Measure of Goodness of Fit for Some Common Nonlinear Regression Models." http://economics.ucdavis.edu/faculty/cameron/research/je97preprint.pdf. Accessed in 2012.

Dissanayake, S. and J.J. Lu (2001). "Factors Influential in Making an Injury Severity Difference to Older Drivers Involved in Fixed Object-passenger Car Crashes." *Accident Analysis and Prevention*, Vol. 34, No. 5, pp. 609-618.

Duncan, C.S., A.J. Khattak, and F.M. Council (1998). "Applying the Ordered Probit Model to Injury Severity in Truck-Passenger Car Rear-End Collisions." *Transportation Research Record*, No. 1635, pp. 63-71.

Eluru, N., M. Bagheri, L.F. Miranda-Moreno, and L. Fu (2012). "A latent class modeling approach for identifying vehicle driver injury severity factors at highway-railway crossings." *Accident Analysis and Prevention*, No. 47, pp. 119-127.

Evans, L. and P.H. Gerrish (2004). "Gender and Age Influence on Fatality Risk from the Same Physical Impact Determined Using Two-car Crashes." *Traffic Safety*. http://www.scienceservingsociety.com/p/137.pdf. Accessed in 2012.

Farr, E.H. (1987). "Summary of the DOT Rail-Highway Crossing Resource Allocation Procedure- Revised." U.S. Department of Transportation, Federal Railroad Administration, Transportation Systems Center, Cambridge, MA. http://www.fra.dot.gov/downloads/safety/grxing_summary.pdf. Accessed in 2012.

Federal Railroad Administration (2010). "Highway-Rail Crossing Program." http://www.fra.dot.gov/rrs/pages/fp_86.shtml. Accessed in 2012.

Federal Railroad Administration Office of Safety Analysis (2012). "Highway-Rail Crossing Inventory Data." http://safetydata.fra.dot.gov/OfficeofSafety/Default.aspx. Accessed in 2012.

Gitelman,V. and A.S. Hakkert (1997). " The Evaluation of Road-rail Crossing Safety With Limited Accident Statistics." *Accident Analysis and Prevention*, Vol. 29, No. 2, pp. 171-179.

Haleem, K., and M. Abdel-Aty (2010). "Examining Traffic Crash Injury Severity at Unsignalized Intersections." *Journal of Safety Research*, Vol. 41, No. 4, pp. 347-357.

Hu, S.R., C.S. Li, and C.K. Lee (2010). "Investigation of Key Factors for Accident Severity at Railroad Grade Crossings by Using a Logit Model." NIH Public Access Author Manuscript, http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2786167/pdf/nihms138304.pdf. Accessed in 2012.

Huang, H., C.H. Chor, and M.M. Haque (2007). "Severity of Driver Injury and Vehicle Damage in Traffic Crashes at Intersections: A Bayesian Hierarchical Analysis." *Accident Analysis and Prevention*, Vol. 40, No. 1, pp. 45-54.

Islam, S. and F. Mannering (2006). "Driver Aging and Its Effect on Male and Female Single-Vehicle Accident Injuries: Some Additional Evidence." *Journal of Safety Research*, Vol. 37, No. 3, pp. 267-276.

Jonsen, S.U. (2007). "Passive Railroad-Highway Grade Crossings." http://www.trafitec.dk/pub/Passive%20railroad-highway%20grade%20crossings.pdf. Accessed in 2007.

Kockelman, K.M. and Y.J. Kweon (2001). "Driver Injury Severity: An Application of Ordered Probit Models." *Accident Analysis and Prevention*, Vol. 34, No. 3, pp. 313-321.

Lam, L.T. (2002). "Distractions and the Risk of Car Crash Injury: the Effect of Drivers' Age." *Journal of Safety Research*, Vol. 33, No.3, pp. 411-419.

Lee, C., and M. Abdel-Aty (2008). "Presence of Passengers: Does It Increase or Reduce Driver's Crash Potential?" *Accident Analysis and Prevention,* Vol. 40, No. 5, pp. 1703-1712.

Lerner, N.D., R.E. Llaneras, H.W. Mcgee, and D.E. Stephens (2002). "Traffic-control Devices for Passive Railroad-highway Grade Crossings" National Cooperative Highway Research Program. http://onlinepubs.trb.org/onlinepubs/nchrp/nchrp_rpt_470-a.pdf. Accessed in 2012.

Massie, D.L. and K.L. Campbell (1993). "Analysis of Accident Rates by Age, Gender, and Time of Day Based on the 1990 Nationwide Personal Transportation Survey." The University of Michigan. http://141.213.232.243/bitstream/2027.42/1007/2/83596.0001.001.pdf. Accessed in 2012.

McCollister, G.M. and Pflaum, C.C. (2007). "A model to predict the probability of highway rail crossing accidents." *Journal of Rail and Rapid Transit,* Vol. 221, No. 3, pp. 321-329.

Meeker, F., D. Fox, and C. Weber (1997). "A Comparison of Driver Behavior at Railroad Grade Crossings with Two Different Protection Systems." *Accident Analysis and Prevention,* Vol. 29, No. 1, pp. 11-16.

Millegan, H., X.D. Yan, S. Richards, and L.D. Han (2010). "Do Stop Signs Improve Safety at Highway-Rail Grade Crossings?" *ITE Journal,* Vol. 80, No. 2, pp. 18-23.

Milton, J., V. Shankar, and F. Mannering (2008). "Highway Accident Severities and the Mixed Logit Model: An Exploratory Empirical Analysis." *Accident Analysis and Prevention,* Vol. 40, No. 1, pp. 260-266.

Miranda-Moreno L.F., L.P. Fu, S. Ukkusuri, and D. Lord (2009). "How to Incorporate Accident Severity and Vehicle Occupancy into the Hotspot Identification Process?" *Transportation Research Record,* No. 2102, pp. 53-60.

Miranda-Moreno, L.F., L.P. Fu, F.F. Saccomanno, and A. Labbe (2005). "Alternative Risk Models for Ranking Locations for Safety Improvement." *Transportation Research Record,* No. 1908, pp. 1-8.

Mohamed, A.A. (2003). "Analysis of Driver Injury Severity Levels at Multiple Locations Using Ordered Probit Models." *Journal of Safety Research,* Vol. 34, No. 5, pp. 597-603.

Morgan, A. and F.L. Mannering (2011). "The Effects of Road-surface Conditions, Age, and Gender on Driver-injury Severities." *Accident Analysis and Prevention,* Vol. 43, No. 5, pp. 1852-1863.

National Transportation Safety Board (1998). "Safety at Passive Grade Crossings.", http://www.ntsb.gov/doclib/reports/1998/SS9802.pdf. Accessed in 2012.

Noyce, D.A. and Daniel B.F. (1998). "Enhanced Traffic Control Devices at Passive Highway-Railroad Grade Crossings." http://www.topslab.wisc.edu/publications/David/noyce_1998_0320.pdf. Accessed in 2012.

Obeng, K. (2011). "Gender Differences in Injury Severity Risks in Crashes at Signalized Intersections." *Accident Analysis and Prevention,* Vol. 43, No. 4, pp. 1521-1531.

Oh, J., Simon P.W., and Doohee N. (2005). "Accident prediction model for railway-highway interfaces." *Accident Analysis and Prevention*, Vol. 38, No. 2, pp. 346-356.

Pai, C.W. and W. Saleh (2007). "An Analysis of Motorcyclist Injury Severity under Various Traffic Control Measures at Three-Legged Junctions in the UK." *Safety Science,* Vol. 45, No. 8, pp. 832-847.

Peck, S.M., Anya A.C., and Miriam K. (2010). "Private Highway-Rail Grade Crossing Safety Research and Inquiry, Volume I.", http://www.fra.dot.gov/downloads/Research/ord1002_II.pdf . Accessed in 2012.

Peterman, D.R., and William J.M. (2009) "The Federal Role in Rail Transit Safety." Members and Committees of Congress. http://assets.opencrs.com/rpts/R40688_20090706.pdf. Accessed in 2012.

Rahimi, M., Najmedin M. (2001). "Human Factors in Highway-Rail Crossing Accidents: the Influence of Driver Decision Style." Proceedings of the First International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design. http://drivingassessment.uiowa.edu/DA2001/70_Rahimi_Mansour.pdf. Accessed in 2012.

Raub, R.A. (2006). "Examination of Highway-Rail Grade Crossing Collisions Over 10 Years." *ITE Journal,* http://www.ite.org/membersonly/itejournal/pdf/2006/JB06DA16.pdf. Accessed in 2012.

Ries, R. (2007). "Highway-Rail Grade Crossing Safety." *Mass Transit*. http://www.masstransitmag.com/article/10221118/highway-rail-grade-crossing-safety. Accessed in 2012.

Rifaat, S., R. Tay, and A. de Barros (2011). "Effect of Street Pattern on the Severity of Crashes Involving Vulnerable Road Users." *Accident Analysis and Prevention,* Vol. 43, No. 1, pp. 276-283.

Ryan, G.A., M. Legge and D. Rosman (1998). "Age Related Changes in Drivers' Crash Risk and Crash Type." *Accident Analysis and Prevention,* Vol. 30, No. 3, pp. 379-387.

Saccomanno,F.F., L.P. Fu, and L.F. Miranda-Moreno (2004). "Risk-Based Model for Identifying Highway-Rail Grade Crossing Blackspots." *Transportation Research Record: Journal of the Transportation Research Board*, No. 1862. pp. 127-135.

Saccomanno, F.F., P.Y. Park, and L.P. Fu (2006). “Estimating Countermeasure Effects for Reducing Collisions at Highway-railway Grade Crossings.”, *Accident Analysis and Prevention*, Vol. 39, No. 2, pp. 406-416.

Savage, I. (2005). “Does public education improve rail-highway crossing safety?” *Accident Analysis and Prevention*, Vol. 38, No. 2, pp. 310-316.

Savolainen, P. and F. Mannering (2007). “Probabilistic Models of Motorcyclists’ Injury Severities in Single-and-multi-vehicle Crashes.” *Accident Analysis and Prevention*, Vol. 39, No. 5, pp. 955-963.

Shankar, V. and F. Mannering (1996). “An Exploratory Multinomial Logit Analysis of Single-Vehicle Motorcycle Accident Severity.” *Journal of Safety Research*, Vol. 27, No. 3, pp. 183-194.

Sposato, S., Patrick B.A., and Mina C. (2006). “Safety of Highway-Rail Grade Crossings.” Federal Railroad Administration. http://www.fra.dot.gov/downloads/research/ord0627.pdf. Accessed in 2012.

Ulfarsson, G.F. and F.L. Mannering (2003). “Differences in Male and Female Injury Severities in Sport-utility Vehicle, Minivan, Pickup and Passenger Car Accidents.”, *Accident Analysis and Prevention*, Vol. 36, No. 2, pp. 135-147.

Woll, T.P. (2007). “The DOT National Crossing Inventory File.”, Eastern Region Highway-rail Crossing Safety Conference Albany, New York. https://www.dot.ny.gov/divisions/operating/osss/rail/grade-crossings/repository/Thomas%20P%20Woll%20-%20FRA%20Inventory%20Update%20ERCC.pdf. Accessed in 2012.

Yan, X.D. (2011) “Using hierarchical tree-based regression model to predict train-vehicle crashes at passive highway-rail grade crossings.” *Accident Analysis and Prevention*, Vol. 42, No. 1, pp. 64-74.

Yan, X.D., L.D. Han, S. Richards, and H. Millegan (2010). “Train-Vehicle Crash Risk Comparison between Before and After Stop Signs Installed at Highway-Rail Grade Crossings.” *Traffic Injury Prevention*, Vol. 11, No. 5, pp. 535-542.

Zhang, J., J. Lindsay, K. Clarke, G. Robbins, and Y. Mao (2000) “Factors Affecting the Severity of Motor Vehicle Traffic Crashes Involving Elderly Drivers in Ontario.” *Accident Analysis and Prevention*, Vol. 32, No. 1, pp. 117-125.

Zhang, Y., Z.B. Li, P. Liu, and L.T. Zha (2011). “Exploring Contributing Factors to Crash Injury Severity at Freeway Diverge Areas Using Ordered Probit Model.” *Procedia Engineering*, Vol. 21, pp. 178-85.

Zwahlen, H.T., and T. Schnell (2000). “Evaluation of the Buckeye Crossbuck at Public, Passive Railroad-highway Grade Crossings in Ohio.” FHWA. http://www.ohio.edu/orite/research/loader.cfm?csModule=security/getfile&pageid=1281538. Accessed in 2012.