An accident prediction model for highway-rail interfaces
by Ross Duane Austin

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in Civil Engineering
Montana State University

Abstract:
Safety levels at railroad/roadway interfaces continue to be of major concern despite an ever-increasing focus on improved design and appurtenance application practices. Despite the encouraging trend toward improved safety, many fatalities continue to occur. Accidents are even happening at public crossings where active warning devices (i.e., gates, lights, bells, etc.) are in place and functioning properly. This phenomenon speaks directly to the need to re-examine both safety evaluation (i.e., accident prediction) methods and design practices at highway-rail crossings.

With respect to safety evaluation methods, the U.S. Department of Transportation's (USDOT) Accident Prediction Formula, is most widely used although three other predominant accident prediction models exist: the Peabody Dimmick Formula, the New Hampshire Index and the National Cooperative Highway Research Program (NCHRP) Hazard Index. Each of these models has strengths, but their shortcomings are apparent.

The Peabody Dimmick Formula, the New Hampshire Index (in its original form) and the NCHRP model are all simple to apply but lack descriptive capabilities due to limited factor considerations. Surprisingly, many similarities exist between the USDOT Accident Prediction Model and the Negative Binomial model developed as part of this investigation with respect to the factors influencing highway-rail crossing accident frequency.

These similarities between the USDOT Accident Prediction Model and the negative binomial model developed here suggest that in fact a successful alternate model has resulted capable of predicting accident frequencies at highway-rail crossings. The benefit to be gained through the development of this alternate model is: (1) a greatly simplified, one-step estimation process, (2) comparable supporting data requirements, and (3) interpretation of both the magnitude and direction of the effect of the factors found to significantly influence highway-rail crossing accident frequencies. However, prior to widespread application of the negative binomial accident prediction model, the model form and estimated coefficients require validation to ensure accuracy in prediction.

AN ACCIDENT PREDICTION MODEL FOR HIGHWAY-RAIL INTERFACES

by

Ross Duane Austin

A thesis submitted in partial fulfillment
of the requirements for the degree

of

Master of Science

in

Civil Engineering

MONTANA STATE UNIVERSITY – BOZEMAN
Bozeman, Montana

July 2000

APPROVAL

of a thesis submitted by

Ross Duane Austin

This thesis has been read by each member of the thesis committee and has been found to be satisfactory regarding content, English usage, format, citations, bibliographic style, and consistency, and is ready for submission to the College of Graduate Studies.

Dr. Jodi L. Carson — Jodi L. Carson (Signature) — 7/13/00 (Date)

Approved for the Department of Civil Engineering

Dr. Donald Rabern — Donald C. Rabern (Signature) — 7/13/00 (Date)

Approved for the College of Graduate Studies

Dr. Bruce McLeod — Bruce R. McLeod (Signature) — 7-18-00 (Date)

STATEMENT OF PERMISSION TO USE

In presenting this thesis in partial fulfillment of the requirements for a master's degree at Montana State University – Bozeman, I agree that the Library shall make it available to borrowers under rules of the Library.

If I have indicated my intention to copyright this thesis by including a copyright notice page, copying is allowable only for scholarly purposes, consistent with "fair use" as prescribed in the U.S. Copyright Law. Requests for permission for extended quotation from or reproduction of this thesis in whole or in parts may be granted only by the copyright holder.

Signature ______________________

Date 7/14/2000

ACKNOWLEDGMENTS

This thesis has been put together with considerable help from several important people. I would like to thank my committee members Jodi Carson, Scott Bell, and John Mounce for all of the time they have put in. Thanks especially to Jodi for the extra time she put in on evenings and weekends to help in the completion of this thesis. Her inspiration, dedication, and statistical expertise were indispensable. An additional acknowledgement must be given to Scott for creating my eternally flexible work schedule regardless of the deadlines involved.

In addition, I would like to thank the Western Transportation Institute for providing the resources necessary to complete this study. I greatly appreciate the opportunities I have been given in order to complete both my education at Montana State University and this thesis.

As always, I am forever indebted to my family and my friends. Through all of the celebrations and struggles, it was your laughter and encouragement that propelled me to finish.

TABLE OF CONTENTS

LIST OF TABLES

Table Page

LIST OF FIGURES

ABSTRACT

Safety levels at railroad/roadway interfaces continue to be of major concern despite an ever-increasing focus on improved design and appurtenance application practices. Despite the encouraging trend toward improved safety, many fatalities continue to occur. Accidents are even happening at public crossings where active warning devices (i.e., gates, lights, bells, etc.) are in place and functioning properly. This phenomenon speaks directly to the need to re-examine both safety evaluation (i.e., accident prediction) methods and design practices at highway-rail crossings.

With respect to safety evaluation methods, the U.S. Department of Transportation's (USDOT) Accident Prediction Formula, is most widely used although three other predominant accident prediction models exist: the Peabody Dimmick Formula, the New Hampshire Index and the National Cooperative Highway Research Program (NCHRP) Hazard Index. Each of these models has strengths, but their shortcomings are apparent.

The Peabody Dimmick Formula, the New Hampshire Index (in its original form) and the NCHRP model are all simple to apply but lack descriptive capabilities due to limited factor considerations. Surprisingly, many similarities exist between the USDOT Accident Prediction Model and the Negative Binomial model developed as part of this investigation with respect to the factors influencing highway-rail crossing accident frequency.

These similarities between the USDOT Accident Prediction Model and the negative binomial model developed here suggest that in fact a successful alternate model has resulted capable of predicting accident frequencies at highway-rail crossings. The benefit to be gained through the development of this alternate model is: (1) a greatly simplified, one-step estimation process, (2) comparable supporting data requirements, and (3) interpretation of both the magnitude and direction of the effect of the factors found to significantly influence highway-rail crossing accident frequencies. However, prior to widespread application of the negative binomial accident prediction model, the model form and estimated coefficients require validation to ensure accuracy in prediction.

CHAPTER 1

INTRODUCTION

Safety levels at railroad/roadway interfaces continue to be of major concern despite an ever-increasing focus on improved design and appurtenance application practices. This Chapter details the magnitude of the problem, provides background information to further clarify the topic and describes this report's content and organization to assist the reader in navigating the document.

Problem Description

"From 1978 to 1993, wide-ranging, multidisciplinary safety improvement efforts sponsored and performed by the Federal Railroad Administration (FRA), in partnership with the various other agencies and industry groups, resulted in a 64 percent reduction in the number of grade crossing accidents" (1). Since 1993, this trend toward improved safety and reduced highway-rail crossing accidents has increased to a 69 percent reduction (2).

Despite the encouraging trend toward improved safety, the interface between railroads and roadways still resulted in 431 fatalities in 1998 alone (2). From an economic standpoint, the impact of these fatalities is significant. Using the cost-per-fatality (highway) estimate of $2.6 million established by the Federal Highway Administration (FHWA), the 431 fatalities associated with highway-rail crossings in

1998 amounted to over $1 billion in economic losses (3). The FHWA's cost-per-fatality estimate includes lost productivity and costs associated with property damage, medical care, insurance, funeral requirements, legal activities and other (3).

Perhaps the most disturbing characteristic of highway-rail crossing accidents is that over 50 percent occurred at public crossings where active warning devices (i.e., gates, lights, bells, etc.) were in place and functioning properly (4). This phenomenon speaks directly to the need to re-examine both safety evaluation (i.e., accident prediction) methods and design practices at highway-rail crossings.

With respect to safety evaluation methods, the U.S. Department of Transportation's (USDOT) Accident Prediction Formula, developed in the early 1980's, is most widely used although other accident prediction models exist. (The USDOT Accident Prediction Formula and other models are discussed in greater detail in Chapter 2 of this report.) While this complex, three-part formula comprehensively addresses characteristics that may influence a crossing's level of safety (i.e., train and traffic volumes, site and surface characteristics, road/rail-side appurtenances, etc.), the formula does not readily provide the magnitude to which each of the characteristics contribute to a crossing's level of safety. This shortcoming makes it difficult to identify or prioritize design or improvement activities that will most effectively address safety-related problems.

Background

A highway-rail crossing consists of both highway and railway components. Highway components include drivers, vehicles, the roadway, and pedestrians. Railway components include train and track elements (5). In addition to the components of the physical system, one must also consider historical legislation that has impacted or is currently impacting highway-rail crossings.

Highway Components

An important element in the highway system is the driver. Driver actions on the highway can occur in one of three zones:

(1) the approach zone,

(2) the non-recovery zone and/or

(3) the hazard zone.

In the approach zone, the driver recognizes a crossing ahead and considers the current conditions. In the non-recovery zone, the driver initiates a stopping maneuver while looking for more information concerning the location of the train. Once the driver is in the hazard zone, they must decide whether or not to proceed through the crossing (5).

A vehicle driver's responsibilities at a highway-rail crossing are defined by the Uniform Vehicle Code (UVC). This Code describes the various actions a driver must take at a crossing.

- "**Approach Speed (Sec. 11-801).** No person shall drive a vehicle at a speed greater than is reasonable and prudent under the conditions and having regard to the actual and potential hazards then existing. Consistent with the foregoing, every person shall drive at a safe and appropriate speed when approaching and crossing an intersection or railroad grade crossing…"
- "**Passing (Sec. 11-306).** No vehicle shall be driven on the left side of the roadway under the following conditions:
  - when approaching within 100 feet of or traversing any…rail highway crossing unless otherwise indicated by official traffic control devices…"
- "**Stopping (Sec. 11-701).** Obedience to signal indicating approach of train. Whenever any person driving a vehicle approaches a rail highway crossing under any of the circumstances stated in this section, the driver of such vehicle shall stop within 50 feet, but not less than 15 feet from the nearest rail of such railroad, and shall not proceed until he can do so safely. The foregoing requirements shall apply when:
  - a clearly visible electric or mechanical signal device gives warning of the train;
  - a crossing gate is lowered or when a human flagman gives or continues to give a signal of the approach or passage of a railroad train;
  - a railroad train approaching within approximately 1,500 feet of the highway crossing emits a signal audible from such distance and such railroad train, by reason of its speed or nearness to such crossing, is an immediate hazard and

– an approaching railroad train is plainly visible and is in hazardous proximity to such crossing" (5).

At highway-rail crossings, different types of vehicles and their related performance characteristics challenge these UVC guidelines. Various vehicle dimensions, braking performance, acceleration performance, and cargoes (i.e., bus passengers, hazardous materials, etc.) must be considered in the design or safety improvement process.

For the roadway approach to highway-rail crossings, many different considerations also exist. These include the following:

- location,
- traffic volumes,
- geometric features,
- number of lanes,
- alignment and sight distances,
- crossing surfaces,
- intersecting highways and
- illumination (5).

Pedestrians are the final consideration with respect to the highway component of highway-rail crossing safety and design. Their movements can be controlled through the use of fences, grade separations, education, enforcement, and additional signing (5).

Railroad Components

Railroad components include trains and tracks. Vast differences in train length, weight, number of engines, number of cars, and travel speeds challenges the accurate provision of safe day-to-day operations.

Tracks are not as variable. Six different classes describe railroad tracks. Class type is determined by maximum train speed allowed (see Table 1). Tracks are further described as main, branch, siding, and industry depending on train activity. Main tracks are used for through movements, while branch tracks typically provide the movement of freight to main lines. Siding and industry tracks are used to store, load, and unload rail cars (5).

This investigation focuses only on the safety effects of the physical road and rail infrastructure and site conditions. Vehicle and driver characteristics are more difficult to correct for and hence are not considered in any detail here.

Table 1. Railroad Track Classification

	Train Speed (MPH)	
Track Class	Passenger	Freight
6	110	110
5	90	80
4	80	60
3	60	40
2	30	25
1	15	10

Legislation

While physical infrastructure, site conditions, vehicle and driver characteristics have a direct effect on the level of safety experienced at highway-rail crossings, legislative activity, which can increase both focus and funding related to highway-rail crossings, is an important consideration for this investigation.

The Highway Safety Acts of 1973 and 1976 and the Surface Transportation Assistance Acts of 1978 and 1982 authorized federal funding to states for the purpose of improving safety at public rail-highway crossings. These Acts also provided money for the installation of active signal devices at the crossings. This spurred the U.S. Department of Transportation (DOT) to develop the DOT Rail-Highway Crossing Resource Allocation Procedure (DOT Procedure) (6). The DOT Procedure, using an accident prediction model and a resource allocation model, determines the "crossing safety improvements that result in the greatest accident reduction benefits based on consideration of predicted accidents at crossings, the costs and effectiveness of safety improvement options, and budget limits" (7).

The DOT Procedure simultaneously aids states in their compliance with the Federal Highway Program Manual (FHPM), which specifies that each state have a priority schedule that is based on:

- the potential reduction in the number and/or severity of accidents;
- the cost of the projects and the resources available;
- the relative hazard of public railroad-highway grade crossings based on a hazard index formula;

- on-site inspections of public crossings;
- the potential danger to large numbers of people at public crossings used on a regular basis by passenger trains, school buses, transit buses, pedestrians, bicyclists, or by trains and/or motor vehicles carrying hazardous materials and
- other criteria as appropriate in each State (5).

This investigation most directly addresses the first and third prioritization schedule requirements above: the potential reduction in accident frequency and severity and the development of a hazard index formula.

Report Purpose and Contents

The findings contained in this report respond to the three-part problem described previously and summarized here.

(1) While showing a positive declining trend, highway-rail crossing accidents continue to result in a high number of fatalities annually and therefore require further investigation beyond the state-of-the-practice.

(2) A high proportion of highway-rail crossing accidents occur at locations where active warning appurtenances are in place suggesting that existing strategies for improving safety are ineffective and require re-examination.

(3) Existing safety evaluation methods (i.e., accident prediction models) do not adequately describe design or other characteristics that are most detrimental to

highway-rail crossing safety thus preventing the prioritization or targeting of safety improvements.

This investigation provides for additional focus on highway-rail crossing safety and suggests an improved accident prediction model that allows for greater interpretation of the factors deemed both beneficial and detrimental to highway-rail crossing safety. Using advanced statistical modeling methods, not only can significant contributing factors be identified but the degree to which these factors affect safety at highway-rail crossings can also be determined. Lastly, this investigation will overcome the disjoint between safety-related findings and design or improvement decisions by actively integrating the two sets of information.

Following this introductory material, Chapter 2 examines literature related to: (1) existing accident prediction models, (2) highway-rail crossing design issues, and (3) traditional and advanced safety improvement strategies for highway-rail crossings. Chapter 3 describes the methodology followed as part of this investigation including data collection, reduction and analysis and accident prediction model development. Chapter 4 provides general descriptive statistics related to highway-rail crossing accident characteristics followed by a description of the accident prediction model results and the relationship between the accident prediction model findings and design and improvement decisions. This report concludes with a summary of finding and a series of suggested recommendations in Chapter 5.

CHAPTER 2

LITERATURE REVIEW

A review of literature related to this investigation focused on: (1) previously developed highway-rail crossing accident prediction models, (2) traditional safety improvement measures at highway-rail crossings and (3) emerging or advanced safety improvement measures. Findings from the literature are detailed below.

Accident Prediction Models

A review of the literature revealed four predominant highway-rail crossing accident prediction models in use:

(1) United States Department of Transportation (USDOT) Accident Prediction Formula,

(2) Peabody Dimmick Formula,

(3) New Hampshire Index and

(4) National Cooperative Highway Research Program (NCHRP) Hazard Index.

In addition to these four, several states have developed their own highway-rail crossing accident prediction formula (5).

USDOT Accident Prediction Formula

The USDOT Accident Prediction Formula, developed in the early 1980's, is most widely used. This complex and comprehensive formula comprises three primary equations:

Equation 1:
$$a = KxEIxDTxMSxMTxHPxHLxHT$$

Equation 2:
$$B = \frac{T_o}{T_o + T}(a) + \frac{T}{T_o + T}\left(\frac{N}{T}\right)$$

Equation 3:
$$A = \{0.7159B\} \quad \text{For Passive Devices}$$
$$A = \{0.5292B\} \quad \text{For Flashing Lights}$$
$$A = \{0.4921B\} \quad \text{For Gates}$$

Equation 1 utilizes data elements contained in the Rail-Highway Crossing Inventory (described in greater detail in Chapter 3). Equation 2 introduces accident history as a factor, and combines this with the accident prediction value, *a*, obtained in Equation 1. Equation 3 introduces a normalizing constant that is multiplied by the value produced in Equation 2 (6). These normalizing constants are updated every two years to reflect changes in observed accident rates (8). Equations 1, 2 and 3 are discussed below.

Equation 1. The factors in Equation 1 each represents characteristics of crossings in the Rail-Highway Crossing Inventory (see Tables 2, 3 and 4). These factors were found to be statistically significant, using nonlinear multiple regression, in the prediction of accidents at highway-rail crossings. Notice some important characteristics, such as sight distance, are not included in Equation 1; factors such as sight distance are unavailable in the Rail-Highway Crossing Inventory. Using Table 2, the value calculated represents the factor's influence in the prediction of accidents at highway-rail crossings where:

c = number of highway vehicles per day

t = number of trains per day

mt = number of main tracks

d = number of through trains per day during daylight

hp = highway paved (yes = 1 and no = 2.0)

ms = maximum timetable speed in mph

h1 = number of highway lanes

ht = highway type factor (see Tables 3 and 4 below) (6).

Table 2. Variables for Equation 1 (USDOT Accident Prediction Model)

		Coefficient or Relationship		
Variable	Description	Passive Control	Flashing Lights	Gates
K	Formula Constant	0.002268	0.003646	0.001088
EI	Exposure Index Factor	$((ct+0.2)/0.2)^{0.3334}$	$((ct+0.2)/0.2)^{0.2953}$	$((ct+0.2)/0.2)^{0.3116}$
DT	Day Through Trains Factor	$((d+0.2)/0.2)^{0.1336}$	$((d+0.2)/0.2)^{0.0470}$	1.0
MS	Maximum Speed Factor	$e^{0.0077ms}$	1.0	1.0
MT	Main Tracks Factor	$e^{0.2094mt}$	$e^{0.1088mt}$	$e^{0.2912mt}$
HP	Highway Paved Factor	$e^{-0.6160(hp-1)}$	1.0	1.0
HL	Highway Lanes Factor	1.0	$e^{0.1380(hl-1)}$	$e^{0.1036(hl-1)}$
HT	Highway Type Factor	$e^{-0.1000(ht-1)}$	1.0	1.0

Table 3. Rural Highway Type Values (USDOT Accident Prediction Model)

Highway Type: Rural	Highway Type Factor (ht)
Interstate	1
Other principal arterial	2
Minor arterial	3
Major collector	4
Minor collector	5
Local	6

Table 4. Urban Highway Type Values (USDOT Accident Prediction Model)

Highway Type: Urban	Highway Type Factor (ht)
Interstate	1
Other freeway/expressway	2
Other principal arterial	3
Minor arterial	4
Collector	5
Local	6

Equation 2. Equation 2 adjusts the accident prediction value, a, from Equation 1 to reflect the actual accident history at the crossing (6). The variable, N, is the number of observed accidents in T years at the crossing, and T_0 is the formula weighting factor defined as:

$$T_0 = \frac{1.0}{(0.05 + a)}$$

Equation 3. In Equation 3 above, the normalizing coefficients reflect conditions in 1998 (7). These constants were developed to reflect more the recent accident experiences of highway-rail crossings with similar types of warning devices in place.

The derivation of the normalizing coefficients used in Equation 3 requires some additional dialogue. In essence, the USDOT Accident Prediction Formula is calibrated every two years by comparing a sample of the most recent year's predicted accident frequencies to the actual observed accident frequencies occurring over several previous years. "The process of determining the three new "normalizing constants" for 1998 is performed such that the 1997 accident prediction sum of the top 20 percent of the crossings is made to equal the sum of the observed number of accidents that occurred for those same 20 percent of crossings using the accident data for Calendar Years 1992 to 1996 (to predict 1997)" (2).

Table 5 and Figure 1 report both the most recent normalizing coefficients and normalizing coefficients from previous years. Note in both Table 5 and Figure 1 the steady reduction in normalizing coefficients over time, or in other words, the steady decline in accident prediction model accuracy as compared to observed values. For example, consider gated highway-rail crossings. The value predicted by the USDOT Accident Prediction Formula Equations 1 and 2 is reduced by more than half with the normalizing coefficient of 0.4921 to reflect actual observed safety levels.

Table 5. Normalizing Coefficients for Equation 3 (USDOT Accident Prediction Model)

WARNING DEVICE GROUPS	NEW	PRIOR YEARS			
	1998	1992	1990	1988	1986
(1) Passive	0.7159	0.8239	0.9417	0.8778	0.8644
(2) Flashing Lights	0.5292	0.6935	0.8345	0.8013	0.8887
(3) Gates	0.4921	0.6714	0.8901	0.8911	0.8131

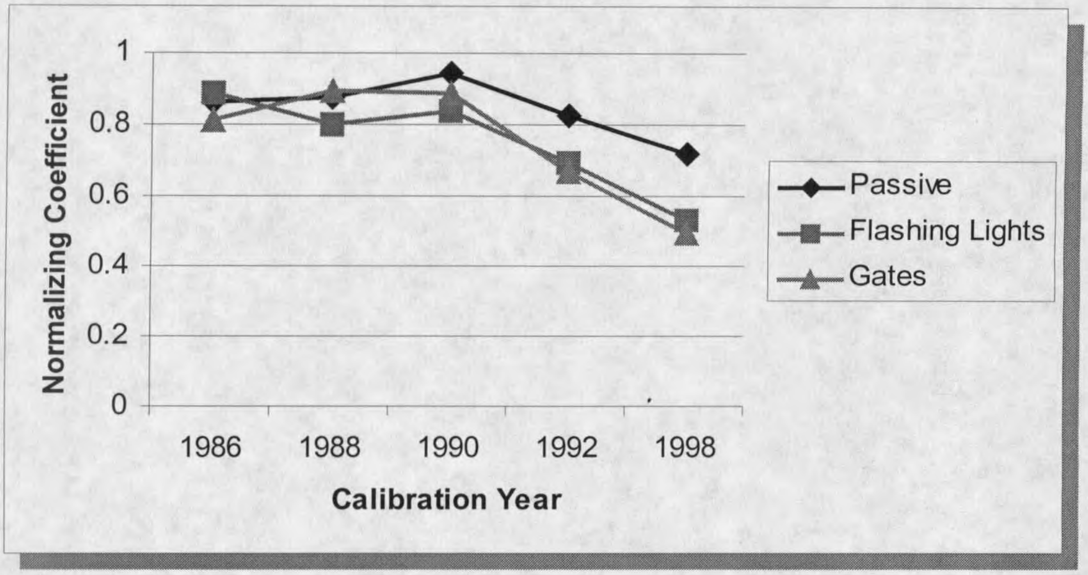

Figure 1. Trend in Normalizing Coefficients for Equation 3 (USDOT Accident Prediction Formula)

Peabody Dimmick Formula

The Peabody Dimmick Formula, also referred to as the Bureau of Public Roads Formula, was developed in 1941 and is used to predict the number of accidents over a five-year time period. This formula was the primary formula utilized from 1941 through the 1950's for resource allocation relating to highway-rail crossings. The specific relationship is as follows:

$$A_5 = 1.28 \frac{\left(V^{0.170} T^{0.151}\right)}{P^{0.171}} + K$$

where:

A_5 = expected number of accidents in 5 years

V = average annual daily traffic (AADT)

T = average daily train traffic

P = protection coefficient (see Table 6)

K = additional parameter (see Figure 2).

In Figure 2 the unbalanced accident factor, I_u, is equal to the first half of the previously listed equation such that:

$$I_u = 1.28\frac{\left(V^{0.170}T^{0.151}\right)}{P^{0.171}}$$

Table 6. Protection Coefficient Values (Peabody Dimmick Formula) (5)

Warning Device	Protection Coefficient P	Warning Device	Protection Coefficient P
Signs	1.65	Wigwag/Flashing Lights/Bells	2.35
Bells	1.78	Watchman, 8 Hours	2.27
Wigwag	1.99	Watchman, 16 Hours	2.43
Wigwag/Bells	2.03	Watchman, 24 Hours	2.52
Flashing Lights	2.18	Gates, 24 Hours	2.56
Flashing Lights/Bells	2.25	Gates, Automatic	2.70
Wigwag/Flashing Lights	2.27		

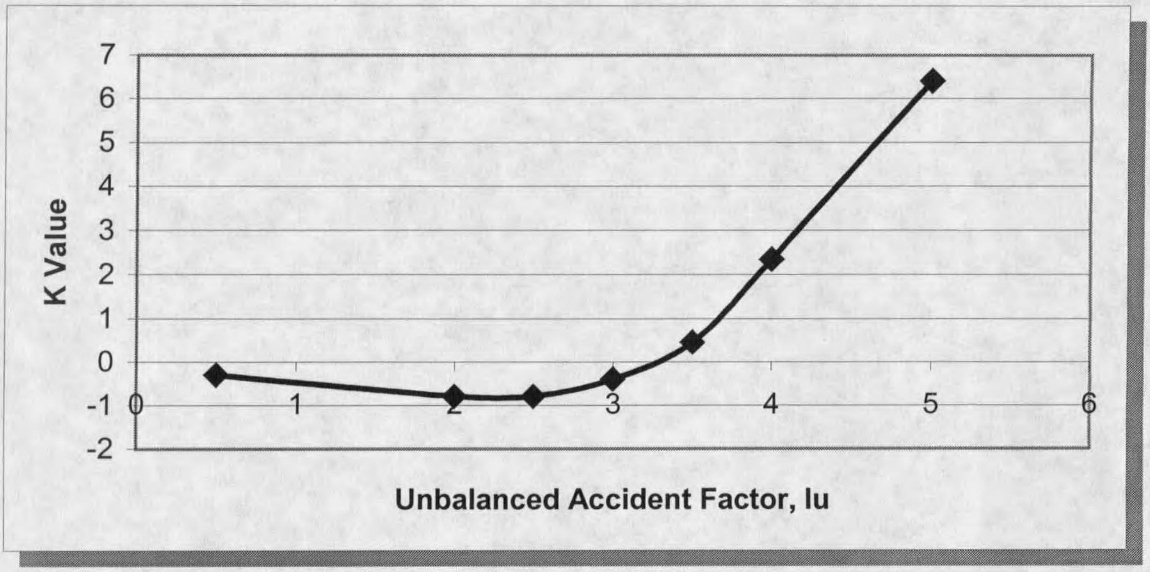

Figure 2. K Values (Peabody Dimmick Formula) (5)

When it was developed, the Peabody Dimmick Formula was based on accident data from rural crossings in 29 states. Non-representative sampling (only rural crossings) hinders the equation's validity. The age of the formula also presents a problem with respect to its ability to predict accidents at crossings where more recent technology is being used.

New Hampshire Index

The New Hampshire Index has a number of variations. The basic formula is:

$$HI = (V)(T)(P_f)$$

where:

HI = hazard index

V = average annual daily traffic (AADT)

T = average daily train traffic

P_f = protection factor.

Several states have developed their own variations of the New Hampshire Index. All utilize the basic equation with modifications to allow incorporation of other accident causative factors. Other factors included in various state formula versions are:

- train speed,
- highway speed,
- sight distance,
- crossing angle,
- crossing width,
- type of tracks,
- surface type,
- number of buses,
- number of passengers,
- and the number of accidents,
- number of tracks,
- nearby intersection,
- functional class of highway,
- vertical alignment,
- horizontal alignment,
- population,
- number of hazardous material trucks and
- number of school buses

Some states also vary the protection factor values:

automatic gates	0.13 or 0.10
flashing lights	0.33, 0.20 or 0.60
wigwags	0.67
traffic signal preemption	0.50
crossbucks	1.00.

Variations of the New Hampshire Index follow:

Variation 1:
$$HI = (V)(2T_f)(T_s)\frac{(SD + AN + NTR)}{4}$$

Variation 2:
$$HI = \frac{(V)(T)(A_5)}{P_f}$$

Variation 3:
$$HI = (V)(T)\frac{(TT + TTR + SD + AN + AL + L + G + VSD + W + LI)}{100}$$

Variation 4:
$$HI = \frac{(P_f)(V_f)(T)(TS)(NTR)}{160} + (70A_a)^2 + 1.2(SD) \qquad A_a = \left(V + \frac{SBP}{1.2}\right)(HM)$$

Variation 5:
$$HI = 0.1(P_f)(A_f)(T_l) + (AN)(NTR)(S)(0.5L) + TS\left((FC)(P) + \left(\frac{(V)(T)}{10{,}000}\right) + SB\right)$$

Variation 6:
$$HI = \frac{(V_f)(P_f)(T)}{(TR + TN + T_f + HS + G + SD + AN)}$$

Variation 7:
$$HI = 0.1(V)(T) + 0.1(HS)(TS) + (SD)(AN)(TR)(NTR)(AL) + (A_a^{\,2} + 1)(RF)(LP)(P_f) + (SB)(SBP) + 10(HM)$$

Variation 8:
$$HI = \frac{(T)\sqrt{(V)}}{P_f}$$

where:

A_5 = number of accidents in five years

A_a = number of accidents per year

A_f = accident factor

AL = highway alignment factor

AN = approach angle factor

FC = functional class factor

G = approach grades factor

HI = hazard index

HM = hazardous material vehicles factor

HS = highway speed factor

L = number of lanes factor

LI = local interference factor

LP = local priority factor

NTR = number of tracks factor

P = population factor

P_f = protection factor

RF = rideability factor

S = surface type factor

SB = number of school buses

SBP=number of school bus passengers

SD = sight distance factor

T = average number of trains per day

T_f = number of fast trains

TN = number of night trains factor

TR = number and type of tracks factor

TS = train speeds factor

T_s = number of slow trains

TT = type of train movements factor

TTR = type of tracks factor

V = annual average daily traffic

V_f = annual average daily traffic factor

VSD = vertical sight distance factor

W = crossing width factor

The dissimilarity between the New Hampshire Index model variations raises concerns over its validity. While most of the discrepancies can be attributed to state preferences, concern is raised due to the lack of consistency. Depending on the variation chosen, prediction values vary considerably.

NCHRP Hazard Index

The National Cooperative Highway Research Program (NCHRP) Hazard Index, documented in NCHRP Report 50, was published in 1964 in a joint effort between the American Association of State Highway Officials (AASHO now AASHTO) and the Association of American Railroads (AAR) in response to the disproportionately high number of accidents occurring at highway-rail crossings. The NCHRP Hazard Index used accident data that spanned five years and was collected by the Interstate Commerce Commission, state agencies and others (9). The NCHRP Hazard Index closely resembles the basic formula of the New Hampshire Index described above:

$$EA = (A)(B)(CTD)$$

where:

EA = expected accident frequency (acc/yr)

A = vehicles per day factor (see Table 7)

B = existing devices factor (see Table 8)

CTD = current trains per day.

Table 7. Vehicles Per Day Factor (NCHRP Hazard Index) (5)

Vehicles per Day	A	Vehicles per Day	A
250	0.000347	9000	0.011435
500	0.000694	10000	0.012674
1000	0.001377	12000	0.015012
2000	0.002627	14000	0.017315
3000	0.003981	16000	0.019549
4000	0.005208	18000	0.021736
5000	0.006516	20000	0.023877
6000	0.007720	25000	0.029051
7000	0.009005	30000	0.034757
8000	0.010278		

Table 8. Existing Devices Factor (NCHRP Hazard Index) (5)

Existing Devices		B
A	Crossbucks, highway volume less than 500 per day	3.89
B	Crossbucks, urban	3.06
C	Crossbucks, rural	3.08
D	Stop signs, highways volume less than 500 per day	4.51
E	Stop signs	1.15
F	Wigwags	0.61
G	Flashing lights, urban	0.23
H	Flashing lights, rural	0.93
I	Gates, urban	0.08
J	Gates, rural	0.19

The NCHRP Hazard Index is concise and easy to use. Unfortunately, this is both its virtue and its vice. There are only three variables to calculate which makes it easy to use, but this limits it's descriptive capabilities. In addition, the determination of an urban versus rural crossing is left to interpretation. This is a key point when looking at the different factor values for flashing lights. A difference of 0.7 exists between the factors for urban and rural settings.

Traditional Safety Improvement Measures

When reviewing literature related to traditional highway-rail crossing safety improvements, one document predominates. The Railroad-Highway Grade Crossing Handbook was developed for the Federal Highway Administration to provide "general information on railroad-highway crossings, including characteristics of the crossing environment and users, and the physical and operational improvements for safe and efficient use by both highway and rail traffic" (5). As such, it has become an influential reference in the discussion of rail-highway crossings.

The Handbook describes five broad categories of design activities that improve the safety of highway-rail crossings:

(1) elimination (physical separation),

(2) passive traffic control devices,

(3) active traffic control devices,

(4) site and operational improvements and

(5) surface improvements (5).

Physical Separation

The physical separation of highway from rail provides the highest level of safety since it eliminates all traffic and train interactions. In addition to safety benefits, reduced delay and alleviated crossing maintenance costs are other advantages to at-grade highway-rail crossing elimination. To attain this physical separation, grade separation, highway and railroad relocation, closure, and abandonment are four options (5).

Grade separation – elevating either the highway or rail at the point of crossing - is attractive in areas of high vehicular traffic because it eliminates train-related stopped delay and reduces vehicular travel times. Grade separation also maintains the existing horizontal alignment and crossing location thereby eliminating driver frustration over changes in access. With grade separation, trains are also able to travel at much higher speeds, making them very popular in areas with high-speed passenger trains. The major deterrent associated with grade-separated crossings is the construction cost (5). Upwards of several million dollars is needed to construct over and underpasses that allows for grade separation (10). This high cost is difficult to justify when other safety-related strategies can be implemented for significantly less (anywhere from $8,500 to $1 million) (11).

Highway and railroad relocation can have similar benefits to a grade-separated crossing. Reduced accidents, diminished delay, and higher train and vehicular speeds can be achieved with relocation. Often, tracks can be consolidated in town to eliminate several of the crossings. Depending on the extent of relocation, this alternative may significantly affect land use and development patterns as well as the economic stability of an area (i.e., employment and development opportunities may also relocate). In addition, the costs associated with relocation, particularly new land acquisition and construction, can be very high (5).

Closing a crossing to ensure physical separation of highway and rail should be considered as an alternative for highway-rail crossing improvements, but only in conjunction with a careful review of accessibility. If an access point (i.e. crossing) is eliminated, there must remain sufficient access points across the tracks to ensure that travel times are not substantially increased. If access is compromised and travel times increase, strong public opposition results. More importantly, emergency personnel in fire trucks, police cars, and ambulances, all rely on the shortest routes possible to save lives (5).

Rail line abandonment provides another option for physical separation of highway rail activities. However, this is not an option that can be implemented at will by transportation professionals. Further, abandoned crossings have many safety concerns. Is the line is actually abandoned? One railroad company might discontinue service along a line but when purchased by another company, service might resume along this line. With this potential, paving over or removing the tracks is not an option. If highway-rail crossing warning signs are removed indicating abandonment, there is concern that drivers will become accustomed to the lack of trains. This poses a significant problem if the tracks are reactivated (5).

Passive Traffic Control Devices

The addition of passive traffic control devices at a highway-rail crossing provides an inexpensive safety improvement alternative. Two categories of passive control devices exist for highway-rail crossings: signs and striping. Signs are typically warning or regulatory in nature and include:

- crossbuck and number of tracks,
- advance warning,
- advisory speed,
- stop sign and stop ahead,
- do not stop on tracks,
- exempt,
- turn prohibition and

- no passing (5).

Pavement markings supplement the message relayed to drivers by the signs. Pavement markings have several shortcomings in that they can be covered by snow, wear and become less visible due to traffic use, and may not be very visible when wet. However, they are inexpensive and are used throughout the U.S. While both signs and striping are simple and inexpensive to use, their effectiveness in actually improving the level of safety at a highway-rail crossing may be questionable.

Active Traffic Control Devices

To encourage a higher level of safety at highway-rail crossings, active traffic control devices are often used. Active traffic control devices indicate the actual presence or approach of a train, rather than indicating a continuous warning through passive signs or striping. Active traffic control devices include:

- post-mounted or cantilevered flashing lights,
- automatic gates,
- warning bells,
- active advance warning devices and
- highway traffic signals (5).

Post mounted or cantilevered flashing lights consist of two lights that flash alternately. The intent of flashing lights is to draw additional driver attention to the hazard. Use of flashing lights alone relies heavily on driver judgement to ensure safety at the highway-rail crossing.

Automatic gates provide an added physical barrier to help ensure safety. Gates extend across the oncoming lane of traffic to deter drivers from crossing the tracks while a train is approaching. Three lights typically extend down the length of the gate arm; the one at the end of the gate is solid red, while the other two flash alternately. Flashing lights, typically mounted on the same post as the gate arm, are activated simultaneously to maximize warning effectiveness (5).

Audible warning bells are also frequently used in conjunction with automatic gates and flashing lights. Audible warning bells, typically mounted on top of the same post as the gate arm, are most effective with pedestrians and bicyclists (5).

An active advance-warning device is an advance warning sign that has supplemental yellow hazard lights. They are primarily used when the flashing lights at the crossing are positioned in such a way that the driver cannot react to them quick enough to stop. The hazard lights on an advanced warning sign flash as an early warning if a train is detected. These lights can be located on the top and bottom of the advance warning sign or both lights can be on top of the sign. In addition to the hazard lights, active advance warning signs can have a passive or active message sign attached to them to inform motorists of an approaching train. The passive sign consists of a "train when flashing" message located near the advanced warning sign. Active signs are usually three-piece folding signs that switch from a "stop ahead" message when open, to a "XXX feet" message when closed. The folding system activates at the same time as the flashing lights (5).

Highway traffic signals are utilized when an intersection is in proximity to a highway-rail crossing. When a train is approaching, the standard traffic signal timing is interrupted and adjusted to prevent vehicle queuing at the crossing, and vehicle backup into other nearby intersections (5).

Site and Operational Improvements

Various site and operational improvements are presumed to improve highway-rail crossing safety. These improvements relate to the following:

- safety barriers,
- flagging,
- site distance,
- geometrics,
- illumination and
- miscellaneous (5).

When either passive or active traffic control devices are used, safety barriers should be considered simultaneously to ensure the traffic control devices themselves do not present a hazard. Guardrail can be used to protect both drivers and traffic control devices but caution must be exercised to ensure that the guardrail does not redirect traffic into a greater hazard (i.e., approaching train or oncoming traffic). A situation in which this occurs is rare, but can happen (5).

At crossings where only passive control devices are present or where many switching operations take place, a temporary flagger may be considered for active times.

Site distance considerations include the distance from a vehicle to the crossing, the distance from a vehicle to the tracks, the distance down the tracks that is visible to a vehicle stopped at the crossing, and the approach speed of both vehicular and train traffic. When looking to improve sight distance, a temporary obstruction can be removed at a minimal cost. Permanent structures in the line of sight are more costly to eliminate, but not impossible (5).

Improvements to the horizontal and vertical alignments of the crossing for improved sight distances are usually the most costly to correct (5).

Ideally, the highway should intersect the railroad at right angles and the track should be slightly higher in elevation than the approaches (but not so high as to result in high-centering of vehicles). This provides the driver of the vehicle the best sight distance, and ensures proper drainage away from the crossing thus reducing the potential for ice and rain to accumulate. Curves in both the highway and the rail should be avoided because it takes the driver's attention away from the crossing. Many times, curves are unavoidable due to right-of-way restrictions (5).

In areas where nighttime accidents frequently occur, illumination is often recommended to improve the level of safety. Illumination has been recommended if one or more of the following conditions exist:

- nighttime train operations;
- low train speeds;
- blockage of crossings for long periods at night;
- accident history that indicates that motorists often fail to detect trains at night;

- horizontal and/or vertical alignment of highway approach such that vehicle headlight beam does not fall on the train until the vehicle has passed safe stopping distance;
- long dark trains, such as unit coal trains;
- restricted sight or stopping distance in rural areas;
- humped crossings where oncoming vehicle headlights are visible under train;
- low ambient light levels and
- a highly reliable source of power (5).

Recommendations for the type and location of illumination devices can be found in the Roadway Lighting Handbook that is published by the Federal Highway Administration (FHWA) (5).

Surface Improvements

Crossing surfaces predominately includes asphalt, but timbers and planks are also used. Crossing surfaces are categorized into two different groups: monolithic and sectional. Monolithic surfaces consist of materials like asphalt or poured-in-place concrete. They are formed at the crossing and have to be destroyed upon removal. Conversely, sectional surfaces typically consist of concrete, steel, or rubber that is pre-cast or pre-made off-site and can be easily removed (5).

Advanced Safety Improvement Measures

The preceding section described various traditional safety improvement measures. The effectiveness of these traditional measures in improving highway-rail crossing safety has been questioned. For example, the North Carolina Department of Transportation (NCDOT) recently completed a study in which 42 percent of the vehicle drivers delayed zero seconds before violating a gate closure (12). Because of this suspected ineffectiveness of traditional safety improvement measures, new approaches to crossing safety are being investigated and applied. Some of these advanced measures build upon conventional methods, while others take advantage of Intelligent Transportation Systems (ITS) developments. Advanced safety improvement measures can be generally categorized as the following:

- physical blockage,
- real-time vehicle sensing,
- motorist information and
- enforcement.

Physical Blockage

Physical blockage safety measures include improved automatic gates and median barriers. Traditional gate arms extend across all lanes of traffic in one direction on each side of the crossing. This layout provides a large enough gap between the gates on either side of the crossing for violators to drive through. An improved approach is to use longer gate arms that extend at least ¾ of the way across the entire roadway. These longer gates have proven to be 67 percent effective in deterring motorists from maneuvering around the gates. A typical long-arm gate system costs approximately $8,500 per location (11).

Vicinity power lines can limit use of the aforementioned longer gate arms. When this occurs, articulated gate arms can be employed. Like the longer gate arms, articulated gates still extend at least ¾ of the way across the entire roadway, but the last third of the articulated arm is retractable to allow for better storage and use in height limited areas. The North Carolina Department of Transportation (NCDOT) found these gates to reduce crossing violations by 78 percent (12).

Quadrant gate assemblies incorporate the use of four gate arms at a standard crossing (See Figure 3). Two are placed on each side of the tracks to ensure the intersection is fully blocked. A primary concern with this system is that a vehicle will get trapped on the crossing when the gates lower. Delayed timing, so that the exiting traffic lanes remain open a little longer, may alleviate this concern. Studies have shown that the quadrant gate setup reduces crossing violations by 86 percent. The downside is that the system costs approximately $125,000 per location to install (11).

While the aforementioned modifications to gate design stand to show improvements in crossing safety, the existing crossing infrastructure may limit their effectiveness. Existing signal circuitry and detection systems are not designed for today's faster trains and consequently aren't as reliable as they could be. If the hazard is not accurately identified, the warning devices cannot be deployed properly.

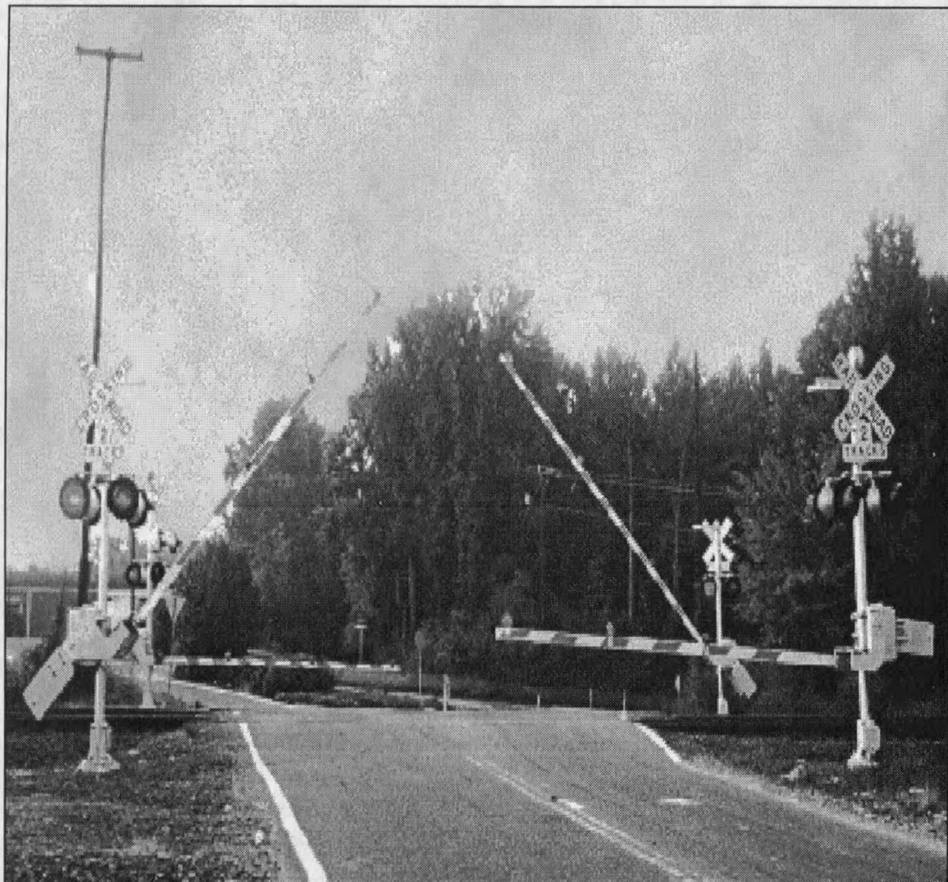

Figure 3. Quadrant Gate Assembly (Photo Provided by North Carolina Department of Transportation)

Unlike gate modifications, median barriers operate passively and therefore aren't subjected to active signal or detection reliability issues. Median barriers in highway-rail crossing applications typically employ a composite speed bump with three-foot high, flexible plastic, reflective paddles to prevent motorists from driving around gates as they close (see Figure 4). Median barriers have shown a 77 percent reduction in violations while costing only $9,100 per location (11).

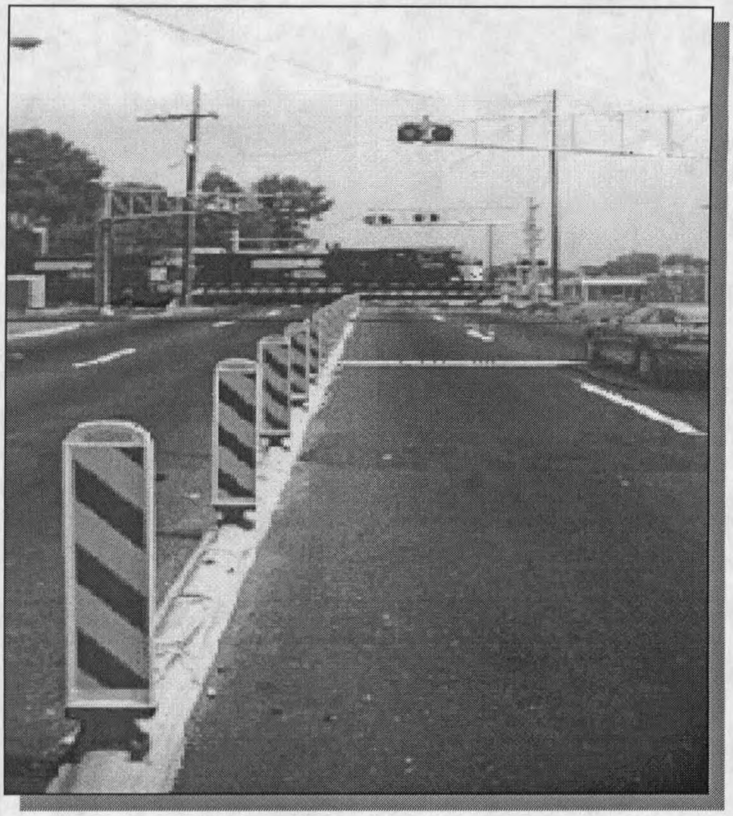

Figure 4. Median Barriers (Photo Provided by North Carolina Department of Transportation)

Barrier gates provide yet another alternative to prevent motorists from driving around gate arms. Barrier gates have many variations but typically consist of some type of wire mesh fence that lowers, blocking an entire road or just one lane of traffic. These formidable gates are attractive in areas where whistle bans are in affect.

Vehicle Arresting Barrier Systems (VAB), or arrestor nets are the most proactive means to prevent crossing violators and improve safety. VABs are currently used at three grade crossings along the Chicago to St. Louis High-speed Rail Corridor. The VAB systems consist of flexible wire that "catches" cars as they try to run the crossing, similar to the ones used on aircraft carriers to stop airplanes. According to testing done at the Texas Transportation Institute (TTI), VABs stopped a fully-loaded (80,000-pound) semi-trailer in 100.5 feet (13).

Real-Time Vehicle Sensing

In the Twin Cities of Minnesota, buses were recently equipped with onboard warning systems as part of a test deployment. These warning systems consist of buzzers and blinking lights to denote the approach of a highway-rail crossing, and the approach of a train. Lights in the buses first blink yellow to denote the crossing, then change to red if a train nears a critical approach range. To prevent potential annoyances from unwarranted blinking lights, several of the buses are equipped with angle detectors which are very useful for drivers whose routes run parallel to train tracks for any period of time. These angle detectors require the bus to approach the crossing at almost a right angle before they activate.

Traditional crossbuck signs are equipped with sensors to detect approaching trains. When a train is detected, a signal is sent to "smart" license plates on the buses. These "smart" license plates, in turn, activate the warning system inside of the bus.

Test results of this system were inconclusive with respect to safety improvements. No significant statistical changes were noted between the control group and the experimental groups. Further, surveys and interviews of bus drivers and train operators commented that the system didn't make them more aware at crossings (14).

Motorist Information

The Second Train Coming (STC) Warning System deployed in Maryland at Timonium Road, utilizes variable message signs and strobe lights to warn motorists that a second train is about to enter the crossing. Many existing rail relay circuits can only identify a single train at a crossing. A loading train will trip the circuit, but because the train is not moving, motorists don't see any danger in crossing. The motorist fails to realize their vision is blocked and a moving train can be hidden behind the other (15). The STC system prevents this occurrence. Once activated, the STC system will continue to be activated until both trains are clear of the crossing. A 90-day analysis of this system showed a reduction in risky behavior at the area (16).

Enforcement

In an attempt to improve highway-rail crossing safety through stricter enforcement, cameras have been placed at crossings (see Figure 5). Reckless drivers are "caught in the act" as they maneuver around gate assemblies. As part of a study sponsored by the Los Angeles County Metropolitan Transportation Authority (LACMTA), a loop detector/video camera system has been deployed (4). When motorists attempt to drive under or around crossing gates, the loop detectors sense this movement and initiates two photos. One photo is of the driver's license plate and the other is a close up of their face. The pictures are stamped with the date, time, vehicle speed, and elapsed time since the active warning signal had been activated (17). The violating motorist subsequently receives a ticket in the mail for the infraction. This study has shown a 92 percent decrease in the number of crossing violations since installation (18).

The NCDOT also studied the use of cameras at highway-rail crossings and found them to be 76 percent effective in reducing violations. As cited in the NCDOT study, the price to supply and install a camera system at one intersection was $100,000 (11).

Video cameras have also been used to allow trains to monitor approaching crossings. In San Antonio, Texas, four trains are currently equipped with video monitors that enable the train operator to see if obstacles (such as a car) are present. Because it takes an average freight train traveling 55mph with 100 rail cars a mile or more to stop, this advanced warning is important (19). This technology is estimated to cost $600,000 (20).

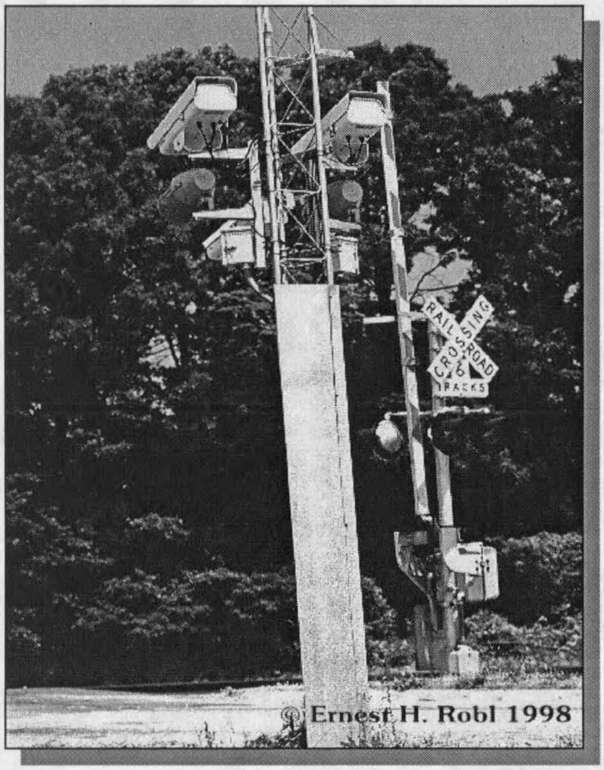

Figure 5. Closed Circuit Television

CHAPTER 3

METHODOLOGY

Data collection procedures and modeling methodologies used to predict the number of accidents at a highway-rail crossing are described in this Chapter. Specifically, this Chapter describes the methodologies related to: (1) data collection and reduction, (2) descriptive statistics, and (3) accident prediction model development.

Data Collection and Reduction

Data to support this investigation came from two sources: (1) the Federal Railroad Administration's Office of Safety accident/incident database and (2) the Federal Railroad Administration's Office of Safety highway-rail crossing inventory. Because of the high number of crossings in the U.S., the data for this investigation was limited to a six-state sample and a two-year time period. The six-state sample was selected by geographical location, number of crossings, and the variability associated with crossings. The states selected for inclusion were California, Montana, Texas, Illinois, Georgia, and New York. In each of these six states, researchers considered highway-rail crossing accidents occurring from January 1997 to December 1998. Only public highway-rail crossings were considered in this investigation.

The FRA's Office of Safety accident/incident database includes accident/incident-specific details about the time, place and conditions of occurrence. Some variables in this database were omitted because of their lack of interest for this investigation. Table 9 lists the accident/incident database elements that were considered. This database was used for two purposes: (1) to provide general descriptive statistics regarding highway-rail crossing accidents and (2) using the grade crossing identification number, to allow calculation of accident frequencies per crossing.

Table 9. Accident/Incident Database Elements

VARIABLE NAME	DEFINITION
GXID	Grade crossing id number
YEAR	Year of incident
MONTH	Month of incident
VEHSPD	Highway user estimated speed
TYPVEH	Highway user type of vehicle
VEHDIR	Highway user direction
POSITION	Position of highway user
TYPACC	Circumstances of accident
TEMP	Temperature in degrees Fahrenheit
VISIBLTY	Daylight period
WEATHER	Weather conditions
TRNSPD	Speed of train in miles per hour
LOCWARN	Location of warning
MOTORIST	Action of motorist
VIEW	Primary obstruction of track view
TOTKLD	Total killed for railroad as reported on f6180.57
TOTINJ	Total injured for railroad as reported on f6180.57
WHISBAN	Whistle ban in effect
DRIVAGE	Vehicle driver's age
DRIVGEN	Vehicle driver's gender

The second database, FRA's Office of Safety highway-rail crossing inventory, provided detailed information about the site and traffic conditions at each of the crossings in the six-state sample. Again, some variables were omitted because of their lack of interest for this investigation. Table 10 lists the variables considered in this investigation.

Table 10. Highway-rail Crossing Inventory Database Elements

VARIABLE	DEFINITION
CROSSING	Highway-rail crossing identification number
DAYTHRU	Average # of daily through train movements between 6 am and 6 PM
NGHTTHRU	Average # of daily through train movements between 6 PM and 6 am
MAXTTSPD	Max timetable speed at the crossing
MAXSPD	Max typical speed at the crossing
MAINTRK	Number of main tracks at the crossing
XBUCK	Number of crossbucks (reflective and non-reflective) at the crossing
STOP	Number of stop signs (standard and non-standard) at the crossing
OTHSGN	Number of signs at the crossing that aren't in any other category
GATE	Number of gates (reflective and non-reflective) at the crossing
FLASH	Number of flashing lights (of all types) at the crossing
HWYSGNL	Number of train activated red-amber-green signals that control street traffic
WIGWAG	Number of wigwags at the crossing
BELLS	Number of bells at the highway-rail crossing, including warning device bells
DEVELTYP	Indicates the predominant type of development in the vicinity of the crossing
HWYPVED	A flag to indicate whether or not the crossing is paved
PAVEMRK	Indicates the presence of highway pavement markings at the crossing
HWYNEAR	Indicates highway intersection by another street, and distance to the intersection
ADVWARN	Indicates advance warning sign presence on any highway approach to the crossing
XANGLE	Classification according to the smallest angle between the highway and the track
XSURFACE	Code for the type of surface covering the crossing
TRAFFICLN	Number of traffic lanes crossing the tracks.
AADT	Average daily traffic over the crossing
PCTTRUK	Percentage of trucks in the traffic stream
DOTACPD	DOT Accident Prediction Value
WHISTBAN	Indicates if a whistle ban is in effect at the crossing
FOURQUAD	Indicates whether or not four-quadrant gates are present
TWOQUAD	Indicates whether or not two-quadrant gates are present
AADTYEAR	The year of the last AADT update
TRAINDAT	The year of the last trains update

Using the FRA's Office of Safety highway-rail crossing inventory directly and supplementing the accident frequency information from FRA's accident/incident database, a combined database was created that formed the basis of the accident prediction model. Information pertaining to the primary sight obstruction at the time of the accident was also included from the first database as this condition was thought to be: (1) important in determining accident frequency, (2) more permanent in nature and (3) not accident-dependent. The variables included in this combined database are listed in Table 11.

Descriptive Statistics

Prior to accident prediction model development, general descriptive statistics describing highway-rail crossing accident characteristics were examined. The FRA's Office of Safety accident/incident database was utilized for this examination. Note that these descriptive statistics represent only the subset of data used for this investigation (i.e., highway-rail crossings and accidents in California, Montana, Texas, Illinois, Georgia and New York for the years 1997 and 1998) and hence, may not reflect national trends. The descriptive statistics considered severity, temporal, environmental, site, vehicle and driver characteristics. Descriptive findings were expressed using histograms, continuous data plots, and pie charts to most clearly display the general trends of highway-rail crossing accidents in the selected states. Findings are detailed in Chapter 4.

Table 11. Combined Database Elements

VARIABLE	DEFINITION
CROSSING	Highway-rail crossing identification number
DAYTHRU	Average # of daily through train movements between 6 am and 6 PM
NGHTTHRU	Average # of daily through train movements between 6 PM and 6 am
MAXTTSPD	Max timetable speed at the crossing
MAXSPD	Max typical speed at the crossing
MAINTRK	Number of main tracks at the crossing
XBUCK	Number of crossbucks (reflective and non-reflective) at the crossing
STOP	Number of stop signs (standard and non-standard) at the crossing
OTHSGN	Number of signs at the crossing that aren't in any other category
GATE	Number of gates (reflective and non-reflective) at the crossing
FLASH	Number of flashing lights (of all types) at the crossing
HWYSGNL	Number of train activated traffic signals that control traffic over the crossing
WIGWAG	Number of wigwags at the crossing
BELLS	Number of bells at the highway-rail crossing including warning device bells
DEVELTYP	Indicates the predominant type of development in the vicinity of the crossing
HWYPVED	A flag to indicate whether or not the crossing is paved
PAVEMRK	Indicates the presence of highway pavement markings at the crossing
HWYNEAR	Indicates highway intersection by another street, and distance to the intersection
ADVWARN	Indicates advance warning sign presence on any highway approach to the crossing
XANGLE	Classification according to the smallest angle between the highway and the track
XSURFACE	Code for the type of surface covering the crossing
TRAFFICLN	Number of traffic lanes crossing the tracks.
AADT	Average daily traffic over the crossing
PCTTRUK	Percentage of trucks in the traffic stream
DOTACPD	Dot accident prediction value
WHISTBAN	Indicates if a whistle ban is in effect at the crossing
FOURQUAD	Indicates whether or not four-quadrant gates are present
TWOQUAD	Indicates whether or not two-quadrant gates are present
AADTYEAR	The year of the last AADT update
TRAINDAT	The year of the last trains update
VIEW	Primary Obstruction of Track View
ACC FREQ	Number of accidents per year

Accident Prediction Model Development

The accident prediction model developed as part of this investigation utilizes the combined FRA databases related to accidents/incidents and highway-rail crossing inventory information. The intent was to develop a model capable of predicting the frequency of highway-rail crossing accidents on the basis of various site and traffic (highway and rail) conditions.

Before performing the modeling exercises, minor data transformations were necessary. Data that had multiple non-numeric choices (i.e., development type = industrial, commercial, residential, etc.) were transformed into singular indicator variables. Data that had multiple choices with a range of numerical values (distance to highway intersection = less than 200 feet, 200 to 400 feet, etc.) were also transformed into indicator variables. Continuous, numeric data, such as the average annual daily traffic, maximum train speeds or percent trucks in the traffic stream were used directly without transformation.

Variables representing the different types of warning devices in place at a crossing could not be used directly because of their endogenous relationship with accident frequency. Note that the relationship of interest is the effect of warning devices on accident frequency. The hypothesis is that warning devices reduce accident frequency. However, it is known that warning devices are placed *in response* to high accident rates resulting in an inappropriate relationship between the dependent variable (i.e., accident frequency) and the independent variable (i.e., the warning device). Should the frequency model be estimated using the warning device indicator variable without correcting for endogeneity, the estimated coefficients would be biased and inconsistent.

To correct for this endogeneity, the method of Instrumental Variables was employed. Each warning device indicator variable was treated as a discrete, binary dependent variable and regressed against all exogenous variables (i.e., site and traffic conditions) using logistic regression:

$$\text{Prob}(i)= e^{\lambda i}/(1+e^{\lambda i})$$

where

$$\lambda_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots \beta_k x_k$$

Using the binary logit model results, warning device probabilities (i.e., a value between 0 and 1) were calculated for every crossing. These newly created variables, one for each type of warning device, were used in the later accident frequency model analysis. The shortcoming of this method is that it is econometrically inefficient.

Once the data transformations were complete, selection of the appropriate model form for the accident prediction model was necessary. A review of prior accident frequency modeling efforts was conducted to help determine the appropriate model form.

Much of the work performed to date has attempted to relate accident frequency to various roadway, environmental, traffic, driver or safety mitigation measure characteristics (see Table 12). Certainly, the bulk of this effort has been put toward relating accident frequency to roadway geometrics and design. Much of the work has considered the highway environment (21, 22, 23, 24). Wong and Nicholson explored the

Table 12. Previous Literature Exploring Accident Frequency (28)

ACCIDENT FREQUENCIES AND GEOMETRIC DESIGN		
Boughton	Accidents and Geometric Design	1975
NCHRP	Cost and Safety Effectiveness of Highway Design Elements	1978
FHWA	Synthesis of Safety Research Related to Traffic Control and Roadway Elements	1982
Miaou, et al	Relationship Between Truck Accidents and Highway Geometric Design: a Poisson Regression Approach	1992
Miaou and Lum	Modeling Vehicle Accidents and Highway Geometric Design Relationships	1993
Shankar, Mannering and Barfield	Effect of Roadway Geometrics and Environmental Factors on Rural Freeway Accident Frequencies	1994
Milton and Mannering	The Relationship Between Highway Geometrics, Traffic Related Elements, and Motor Vehicle Accidents	1996
ACCIDENT FREQUENCIES AND DRIVER BEHAVIOR		
Wong and Nicholson	Driver Behavior at Horizontal Curves: Risk Compensation and the Margin of Safety	1992
ACCIDENT FREQUENCIES AND TRAVEL CHARACTERISTICS		
McGuigan	The Use of Relationships Between Road Accidents and Traffic Flow in Black Spot Identification	1981
Jovanis and Chang	Modeling the Relationship of Accidents to Miles Traveled	1986
ACCIDENT FREQUENCIES AND ENVIRONMENTAL CONDITIONS		
Snyder	Environmental Determinants of Traffic Accidents: An Alternate Model	1974
Ivey, et al	Predicting Wet Weather Accidents	1981
Brodsky and Hakkert	Risk of a Road Accident in Rainy Weather	1988
Shankar, Mannering and Barfield	Effect of Roadway Geometrics and Environmental Factors on Rural Freeway Accident Frequencies	1994
ACCIDENT FREQUENCIES AND COUNTERMEASURE EFFECTIVENESS		
Abbess, Jarret and Wright	Accidents at Black Spots: Estimating the Effectiveness of Remedial Treatment with Special Reference to the 'Regression-to-Mean' Effect	1981
Hanbali	Influence of Winter Road Maintenance on Traffic Accident Rates	1992
ACCIDENT FREQUENCY METHODOLOGIES		
Hammerslag, Roos and Kwakernaak	Analysis of Accidents in Traffic Situations by Means of Multi-proportional Weighted Poisson Model	1982
Maher	Fitting Probability Distributions to Accident Frequency Data	1987
Okamoto and Koshi	A Method to Cope with the Random Errors of Observed Accident Rates in Regression Analysis	1989
Joshua and Garber	Estimating Truck Accident Rate and Involvement Using Linear and Poisson Regression Models	1990
Hauer	Empirical Bayes Approach to the Estimation of Unsafety: The Multivariate Regression Approach	1992

relationship between driver behavior and potentially challenging geometric roadway segments (i.e., horizontal curve sections) (25). Others have attempted to relate traffic flow and vehicle-miles or kilometers traveled to accident frequency (26, 27, 28).

Attempts to analyze accident frequency data have ranged from the use of conventional multiple linear regression using least squares regression techniques to methods involving exponential distributions such as a Poisson (29, 30, 31, 32, 33, 28).

Multiple Linear Regression

Historically, the most common methodological approach used to model the relationship between roadway geometrics, environmental characteristics, traffic characteristics and accident frequency was conventional linear regression.

Model Form. The model form for multiple linear regression is as follows:

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k + \varepsilon$$

where

Y is the dependent variable

β_0 is the estimable regression parameter indicative of the y-intercept value

β_1 through β_k are estimable regression parameters

x_1 through x_k are independent, explanatory variables

ε is the random error or disturbance term (34).

The method of Ordinary Least Squares (OLS) is used to determine estimates of β_0 through β_k. Ordinary Least Squares provides the best fitting plane (line, in simple linear regression) by minimizing the sum of squares of the distances between the observed or actual values, Y_i, and the corresponding predicted values, y_i where $y_i = b_0 + b_1x_{1i} + b_2x_{2i} + \ldots b_kx_{ki}$. Hence, to employ the method of OLS, the sum of the squared deviations must be determined.

$$Q = \Sigma\,(Y_i - y_i)^2 = \Sigma\,(Yi - b_0 + b_1x_{1i} + b_2x_{2i} + \ldots b_kx_{ki})^2$$

where

Y_i is the observed or actual dependent variable value for sample observation, i

y_i is the predicted dependent variable value for sample observation, i

b_0 is the estimated regression parameter indicative of the y-intercept value

b_1 through b_k are estimated regression parameters

x_{1i} through x_{ki} are independent, explanatory variables for sample observation, i (34).

According to the method of OLS, the best estimates of β_0 and β_1 through β_k are those values that minimize the sum of squares for a given sample observation. In other words, the best fitting model will have the smallest deviations between the observed or actual Y_i and the predicted y_i (34).

Gauss and Markov theorized several important characteristics of the resulting OLS estimates. The Gauss-Markov Theorem states that: "The least squares estimators, b_0, b_1, b_2,...b_k are unbiased and have minimum variance among all unbiased linear estimators." Being unbiased implies that the expected value of each estimate is equal to the true value of the parameter.

$E[b_0] = \beta_0$

$E[b_1] = \beta_1$

:

$E[b_k] = \beta_k$

If the estimates were to overestimate or underestimate the parameter systematically, this assumption of unbiasedness would not hold. Minimum variance implies that the sampling distributions of b_0, b_1, b_2, ... b_k have smaller variability than the sampling distributions of any other estimates. Thus, the OLS estimates are more precise that any other estimate (34).

To prove that Ordinary Least Squares estimates are, in fact, the best, linear, unbiased estimates (BLUE), several assumptions must be made. Violation of any one of these assumptions may jeopardize the quality of the resulting estimates.

Normality. First, the random error or disturbance term, ε_i,, is assumed to be normally distributed with mean zero and variance, σ^2.

ε is $N(0, \sigma^2)$

If this assumption of normality is, in fact, false, (as noted through a plot of the residuals) the estimated coefficients, b_0, b_1, b_2...b_k, will be consistent but inefficient. Further, problems arise in hypothesis testing if the normality assumption is violated (34).

Zero Mean. A second assumption required to prove that OLS estimates are the best, linear unbiased estimates, is that the mean of the random error or disturbance term is equal to zero.

$$E[\varepsilon_i] = 0$$

In most cases, the negative and positive error terms will cancel, resulting in a mean of zero. However, measurement errors that positively or negatively skew the data will nullify this assumption. The effects of this violation of the resulting estimates depend on the model form: (1) if the model includes the estimable regression parameter, β_0, the estimate, b_0 is biased but the estimates, b_1, b_2, ... b_k are unbiased and efficient; (2) if β_0 is omitted from the model, the estimates, b_1, b_2,...b_k are biased and inconsistent (34).

Homoscedasticity. Third, the random error or disturbance terms are assumed to be homoscedastic (i.e., have the same variance).

$$E[\varepsilon_i^2] = \sigma^2$$

The assumption of homoscedasticity is often challenged in aggregated data where error variances may exhibit systematic increasing or decreasing trends. Heteroscedasticity problems result in estimates that are still unbiased and consistent but inefficient. Further, this violation invalidates hypothesis tests because confidence intervals are erroneous thus making it difficult to assess the significance of estimated coefficients (34).

Serial Independence. A fourth assumption to prove that the OLS estimates are the best, linear, unbiased estimates is that the random error or disturbance terms are not serially correlated.

$E[\varepsilon_i\varepsilon_j] = 0$ if and only if i does not equal j

If the random errors are serially correlated, the estimated coefficients are still unbiased but may be inefficient. Serial correlation can be easily detected with the estimation of the Durbin-Watson statistic but most often results in temporally or spatially averaged data (34).

Non-stochastic x. Lastly, it is assumed that the independent variable, x, is not random and has fixed values in repeated samples.

$\mathrm{cov}[x_i\varepsilon_i] = 0$

When an independent variable, x, and the error term, ε, are correlated, this assumption is violated. Measurement errors in the data often lead to violation of this assumption. If Y is measured with error, the parameter estimates will be unbiased and consistent. If x is measured with error, the parameter estimates will be biased and inconsistent. If both Y and x are measured with error, the resulting estimates will be biased and consistent (34).

Shortcomings. Conventional linear regression is inappropriate for modeling most count (i.e., accident frequency) data because the model form is not restrained from predicting negative values. With respect to accident frequency, the prediction of a negative value will bias the estimated coefficients. Secondly, heteroscedasticity problems have been noted with the use of conventional linear regression to model accident frequency data. A model that suffers from heteroscedasticity violates the third fundamental Ordinary Least Squares assumption required to prove that OLS are the best, linear, unbiased estimates, namely that disturbance terms have equal variance (28).

Many have documented the inappropriateness of using conventional multiple linear regression to model the relationship between roadway geometrics, environmental characteristics, traffic characteristics and accident frequency (27, 31, 23, 28).

Jovanis and Chang found a number of problems with the use of linear regression to model the relationship between vehicle-miles traveled and accident frequency (27). They discovered that as vehicle-miles traveled increased, so did the variance of the accident frequency. This occurrence directly violated the homoscedasticity assumption of linear regression. As described previously, the effect of this violation invalidates hypothesis tests because confidence intervals are erroneous thus making it difficult to assess the significance of estimated coefficients. Jovanis and Chang concluded that if the objective of their study was to determine the influence of a specific predictor variable (i.e., vehicle-miles traveled) on accident frequency, the failure to properly test for coefficient significance is a serious flaw (28).

In a study conducted by Miaou and Lum, the statistical properties of conventional linear regression models applied to count data were investigated (23). Miaou and Lum concluded that the conventional linear regression models lack the distributional properties to describe random, discrete, non-negative and typically sporadic vehicle accidents on the road. More importantly, when using linear regression, the authors stated that there was no assurance that the expected total number of accidents predicted would be close to the observed totals. Therefore, the use of linear regression models was inappropriate for making probabilistic statements about the occurrences of vehicle accidents on the road (28).

Poisson Regression

In light of the problems associated with linear regression, many turned to Poisson regression as a means to better predict accident frequency.

Model Form. The Poisson regression model specifies that each Y_i is drawn from a Poisson distribution with parameter λ_i, which is related to the independent or explanatory variables, x_1, x_2,...x_k. The probability distribution for Poisson regression is as follows:

$$P(Y_i = y_i) = \frac{e^{(-\lambda_i)}\lambda_i^{y_i}}{y_i!}$$

where

$P(Y_i = y_i)$ is the probability of y occurring for sample observation, i

λ_i is the Poisson parameter which is equal to the expected value of y_i, $E[y_i]$

y_i is the predicted dependent variable value for sample observation, i (34).

To introduce the estimable parameters, the Poisson parameter can be expressed as the following:

$$\lambda_i = e^{(\beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_k x_{ki})}$$

where

λ_i is the Poisson parameter which is equal to the expected value of y_i, $E[y_i]$

β_0 is the estimable regression parameter indicative of the y-intercept value

β_1 through β_k are estimable regression parameters

x_{1i} through x_{ki} are independent, explanatory variables (34).

The more common formulation for the Poisson parameter, λ_i, is the log-linear model:

$$\text{Log } \lambda_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \ldots + \beta_k x_{ki}$$

The estimable parameters in the Poisson model are estimated using standard maximum likelihood (ML) methods. The underlying principle behind ML is that different statistical populations generate different samples; any one sample is more likely to come from some populations rather than others. The method of ML requires that a likelihood function be specified that depends both on the observed dependent variable, Y_i, and on a set of unknown parameters (34).

$$L(\beta) = \prod \frac{[(-e^{\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k})(e^{\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k})]^{n_i}}{n_i!}$$

In general, while maximum likelihood estimates are consistent, they are not unbiased or efficient (34).

One important property of the Poisson distribution is that it restricts the mean and variance of the distribution to be equal.

$$E[Y_i] = Var[Y_i]$$

If this equality does not hold, the data are said to be either underdispersed or overdispersed (34).

$E[Y_i] > Var[Y_i]$ underdispersed

$E[Y_i] < Var[Y_i]$ overdispersed

If the mean/variance equality is violated, the resulting parameter estimates will be biased (34).

To determine whether or not count data is overdispersed and hence, whether or not Poisson regression is appropriate, three approaches can be taken: (1) a regression-based test, (2) a conditional moment test, or (3) a Lagrange multiplier test.

Cameron and Trivedi recommended the following regression-based approach for detecting overdispersion in count data (35). Cameron and Trivedi proposed the following null and alternate hypotheses:

H_0: $\mathrm{Var}[Y_i] = E[Y_i]$

H_a: $\mathrm{Var}[Y_i]=E[Y_i] + \alpha g(E[Y_i])$

Using simple least squares regression as their basis, the authors suggest regressing z_i on w_i where

$$z_i = \frac{[(Y_i - E[Y_i])^2 - Y_i]}{\sqrt{2}E[Y_i]}$$

and

$$w_i = \frac{g(E[Y_i])}{\sqrt{2}E[Y_i]}$$

Using t-statistics to denote significance, if $t^* \leq t_{(1-\alpha/2,\, n-1)}$, the null hypothesis is true, the data is not overdispersed and Poisson is an appropriate model form. If $t^* > t_{(1-\alpha/2,\, n-1)}$, the alternative hypothesis is accepted, the data are overdispersed and Poisson is an inappropriate model form to use.

The conditional moment method of testing for overdispersion in count data provides a more general hypothesis than the previous regression-based test in that the form of the overdispersion is not specified. The null and alternate hypothesis for the conditional moment test of overdispersion are as follows:

H_0: Var $[Y_i] = \lambda_i$

H_a: $E[Y_i] \neq \lambda_i$

The first derivatives and moment restrictions are as follows:

$E[x_i(Y_i - \lambda_i)] = 0$

$E[z_i((Y_i - \lambda_i)^2 - \lambda_i)] = 0$

To perform the conditional moment test, the following steps are required. Letting $e_i = y_i - \lambda_i^*$ and $z_i = x_i$ with no constant term, compute the Poisson regression. Compute

$$r = \sum z_i[e_i^2 - \lambda_i^*] = \sum z_i v_i$$

using the maximum likelihood estimates from the Poisson regression. Compute the Chi-squared test statistic, χ^{2*}

$$\chi^{2*} = r'S^{-1}r$$

where

$$S = M'M - M'D(D'D)^{-1}D'M$$

and

$$M'M = \sum z_i z_i' v_i^2$$

$$D'D = \sum x_i x_i' e_i^2$$

and

$$M'D = \sum z_i x_i' v_i e_i$$

If $\chi^2_{(\alpha/2,\ n-1)} < \chi^{2*} \leq \chi^2_{(\alpha/2,\ n-1)}$, the null hypothesis is true. Otherwise the alternate hypothesis is true (34).

The Lagrange multiplier test can be used if Poisson model and the alternative model form, the negative binomial model, can be viewed as restricted and non-restricted models, respectively. Let

H_0: $c(\theta) = 0$

H_a: $c(\theta) \neq 0$

where H_0 represents the restricted model state of interest. The Lagrange Multiplier test statistic is as follows:

$$LM = \frac{\sum w_i[(y_i - \lambda_i)^2 - y_i]}{\sqrt{2\sum w_i^2 \lambda_i^2}}$$

where the weight, w_i, varies depending on the assumed distribution. For the negative binomial model, $w_i = 1.0$ (34).

The primary advantage to the Lagrange multiplier test is that you only need to estimate the Poisson model to compute it. As with the conditional moment test, the Lagrange Multiplier test follows a Chi-squared distribution for hypothesis testing; if $\chi^2_{(\alpha/2,\,n-1)} < LM \leq \chi^2_{(\alpha/2,\,n-1)}$, the null hypothesis is true. Otherwise the alternate hypothesis is true (34).

While these tests are comparable in their outcomes (although differences may be noted in small sample applications), the regression-based test proposed by Cameron and Trivedi (35) is the favored test in practice.

Shortcomings. Jovanis and Chang supported the notion that the statistical properties of Poisson regression were superior to those of linear regression for applications regarding highway safety (27). Similarly, Joshua and Garber studied the relationship between highway geometrics and truck accidents in Virginia using both linear and Poisson regression models (31). They also concluded that linear regression techniques used in their research did not describe the relationship between truck accidents and the independent variables but that Poisson techniques did (28).

In other work, Miaou et al. used a Poisson regression model to establish the empirical relationship between truck accidents and highway geometrics on a rural interstate in North Carolina (22). The estimated Poisson model suggested that average annual daily traffic per lane, horizontal curvature and vertical grade were significantly correlated with truck accident likelihood. During their work, a limitation of the Poisson model was uncovered - the constraint that the mean and variance of the accident frequency variable are equal (28).

In most accident data, the variance of accident frequency exceeds the mean and in such cases the data are said to be overdispersed. If corrective measures are not taken, the estimable parameters will be biased (28).

The overdispersion uncovered by Miaou and his colleagues was attributable to uncertainties in the data or omitted variables. They argued that although overdispersion was present, it did not change the conclusion about the relationship between truck accidents and the examined traffic and highway geometric design variables. However, they did suggest that other discrete distributions should be explored (28).

Miaou and Lum (23) conducted a follow-up study. While this study was similar in scope to the first, the main purpose was to evaluate the statistical properties of conventional linear regression models and two Poisson regression models. The four types of models considered were (1) additive linear regression, (2) a multiplicative linear regression model, (3) a multiplicative Poisson model with exponential rate function and (4) a multiplicative Poisson regression model with a nonexponential rate function. The authors found that the Poisson regression models outperformed linear regression models. Furthermore, the Poisson model with the exponential rate function was the favored model (28).

Miaou and Lum also attempted to address overdispersion in their frequency data. When overdispersion existed in the data and a Poisson model was used, the variances of the estimated model coefficients tended to be underestimated. Miaou and Lum attempted to relax the Poisson constraint of the mean being equal to the variance by using Wedderburn's overdispersion parameter. They found that with such overdispersed data, using the Poisson model may not be appropriate for making probabilistic statements about vehicle accidents because the model may under- or over-estimate the likelihood of occurrence. Because of overdispersion difficulties, the authors suggested the use of a more general probability distribution such as the negative binomial (28).

Negative Binomial Regression

Use of negative binomial regression techniques was the next evolutionary step in relating accident frequency to various explanatory variables.

Model Form. The negative binomial model is more appropriate for overdispersed data because the model relaxes the constraint of equal mean and variance ($E[Y_i] = Var[Y_i]$). This relaxation of the Poisson constraint is accomplished through the addition of a Gamma-distributed error term to the Poisson model such that

$$\text{Log } \lambda_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \ldots + \beta_k x_{ki} + \xi_i$$

where

λ_i is the Poisson parameter which is equal to the expected value of Y_i, $E[Y_i]$

β_0 is the estimable regression parameter indicative of the y-intercept value

β_1 through β_k are estimable regression parameters

x_1 through x_k are independent, explanatory variables

ξ_i is the Gamma-distributed error term (34).

The addition of ξ_i allows the mean to differ from the variance such that

$$Var[Y_i] = E[Y_i]\,[1 + \gamma E[Y_i]]$$

where γ is an additional estimable parameter. If γ is not significantly different from zero, the data is not overdispersed and the Poisson regression is appropriate (34).

LIMDEP Computer Program

To analyze this data, the computer program LIMDEP was utilized because its capabilities include everything from basic linear regression to Poisson regression models and nonlinear regressions estimated through instrumental variables. Another reason it was chosen was because it can also handle large data sets and many variables (36).

CHAPTER 4

HIGHWAY-RAIL CROSSING SAFETY CHARACTERISTICS

This Chapter describes the factors that significantly influence the level of safety experienced at highway-rail crossings as determined through the development of an accident prediction model. Preceding this discussion, is a summary of descriptive highway-rail crossing accident characteristics.

Accident Characteristics

Prior to accident prediction model development, general descriptive statistics describing highway-rail crossing accident characteristics were examined. Note that these descriptive statistics represent only the subset of data used for this investigation (i.e., highway rail crossings and accidents in California, Montana, Texas, Illinois, Georgia and New York for the years 1997 and 1998) and hence, may not reflect national trends. The descriptive statistics provided below consider severity, temporal, environmental, site, and vehicle and driver characteristics.

Severity

As seen in Figures 6 and 7 below, accidents have severe results at highway-rail crossings. For this sample data, one fatality occurred for every 7.7 accidents. This is slightly higher than the national average at highway-rail crossings over this same time period of one fatality for every 8.3 accidents (37).

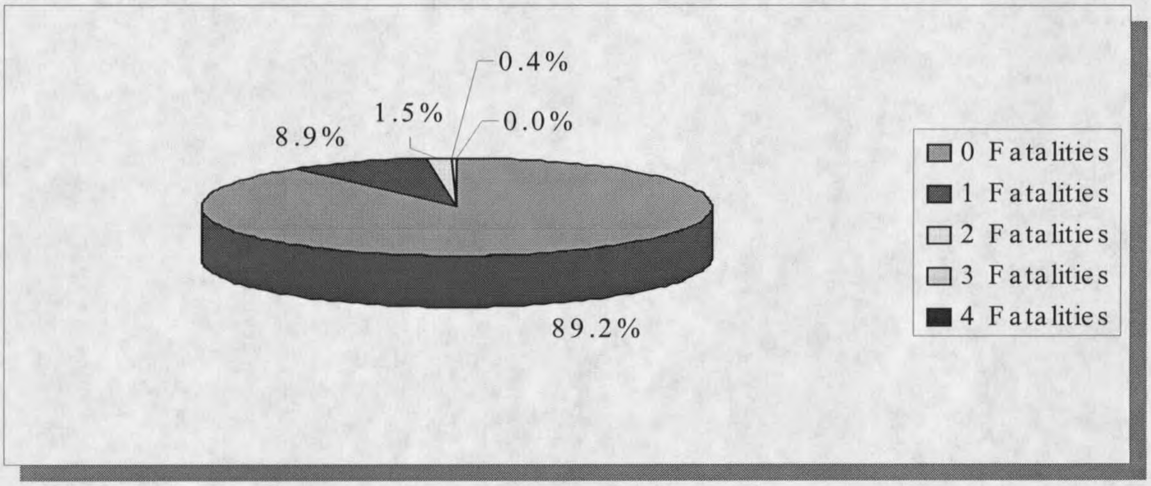

Figure 6. Fatalities per Highway-rail Crossing Accident

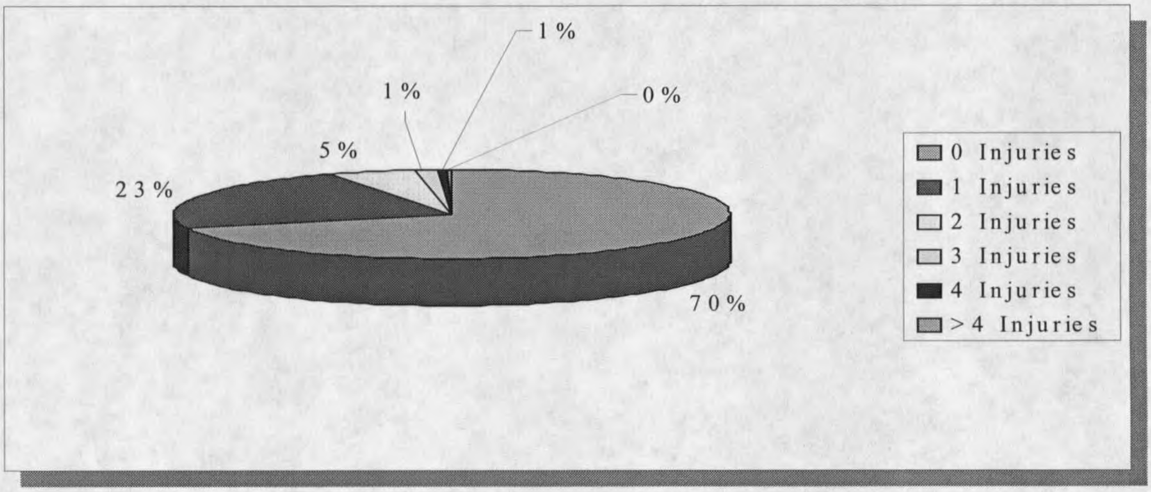

Figure 7. Injuries per Highway-rail Crossing Accident

Temporal Characteristics

On average, 3,687 highway-rail crossing accidents occurred annually over the sample period (i.e., January 1997 to December 1998). Considering accident frequency variation throughout the year, peaks occurred in January, March, July, November, and December with the most noted peak occurring in March (see Figure 8). Why March has the highest number of accidents isn't intuitively obvious.

When looking at the accident frequency variation throughout the day, over twice as many accidents occurred during the daytime as compared to nighttime (64 percent versus 31 percent) (see Figure 9). Larger vehicular traffic volumes during the day may explain this.

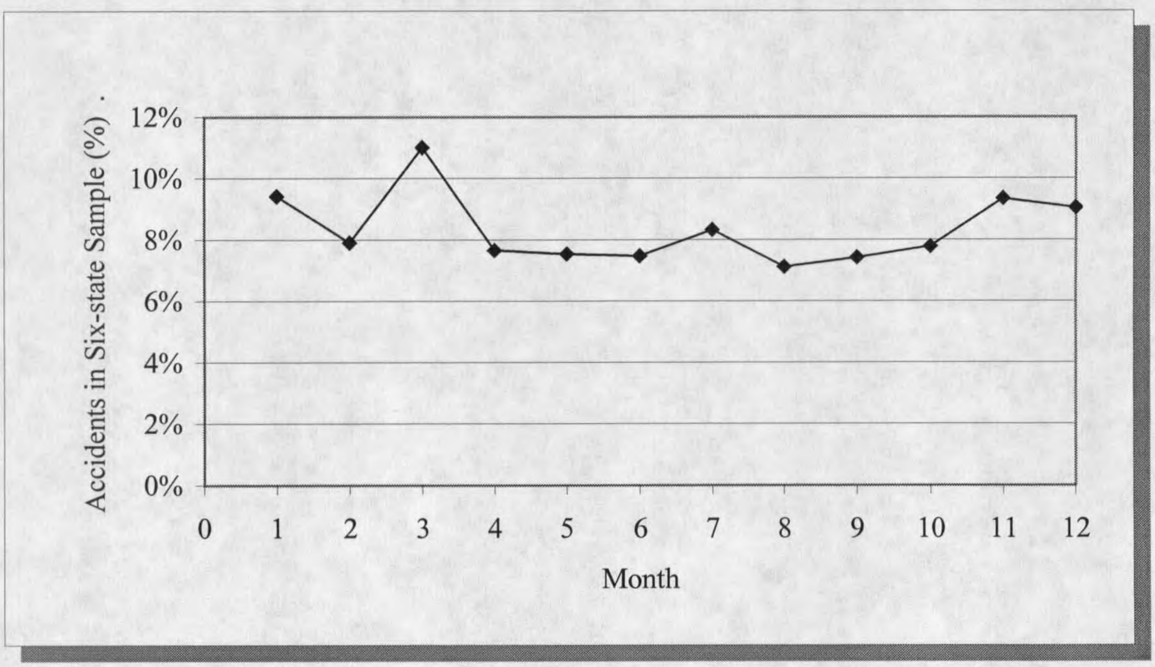

Figure 8. Highway-rail Crossing Accidents by Month

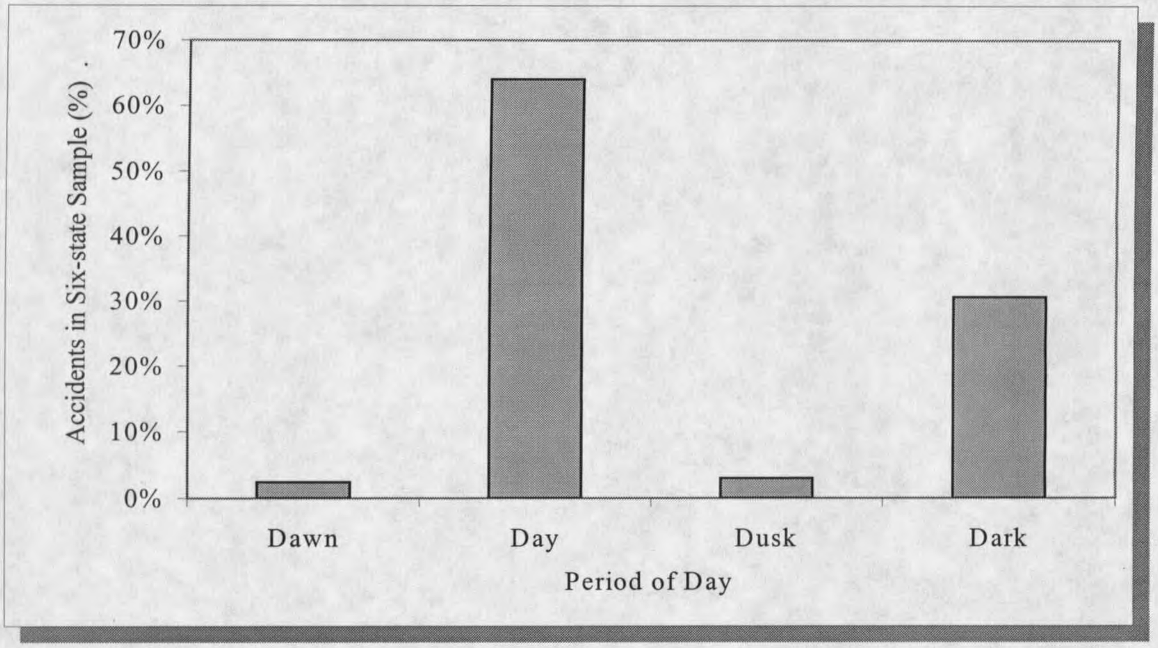

Figure 9. Highway-rail Crossing Accidents by Period of Day

Environmental Characteristics

Weather effects such as snow and ice seem inconsequential with respect to highway-rail crossing accidents. The highest accident frequencies were noted when the outside air temperature averaged 70 degrees (see Figure 10). This is somewhat surprising when one considers the information presented in Figure 8 which showed higher accident frequencies in the winter and spring months (i.e, January, March, November, December). The explanation of this becomes clearer when it is noted that over 55 percent of the accidents in this data set occurred in Texas and Georgia.

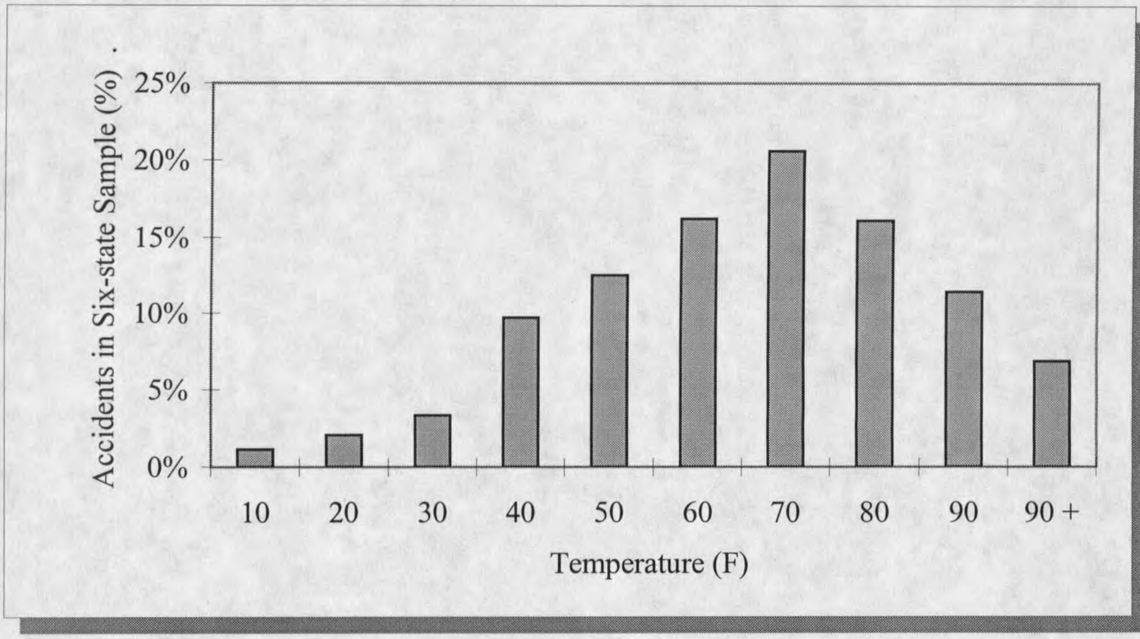

Figure 10. Highway-rail Crossing Accidents by Ambient Temperature

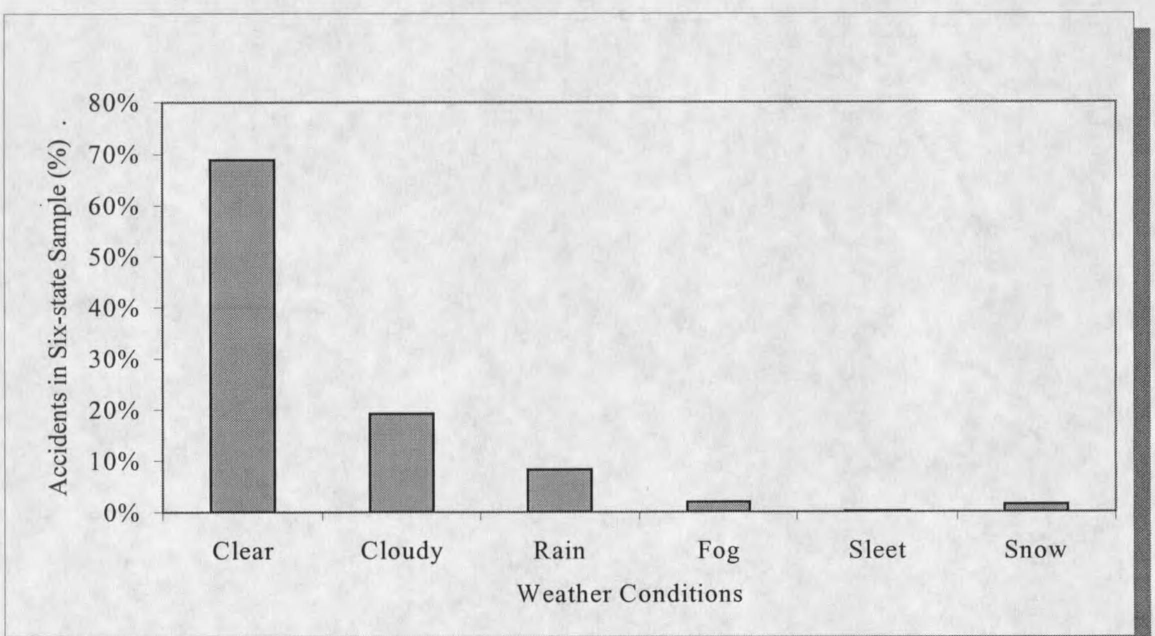

Figure 11. Highway-rail Crossing Accidents by Weather Conditions

Most accidents occurred when the sky was clear (69 percent) or overcast (19 percent) (see Figure 11). Approximately 8 percent of the accidents occurred when it was raining. Fog, sleet, and snow only existed in a combined 4 percent of accidents.

Site Characteristics

With respect to site characteristics, both driver visibility and existing warning systems were examined. With respect to driver visibility, it was hypothesized that a high number of accidents resulted because the motorist's view of the approaching train was obstructed. Disproving this hypothesis, only 10 percent of the crossings had some form of obstacle in the way of the train and/or tracks at the time of the accident (see Figure 12).

Turning attention to warning systems in place at highway-rail crossings, the position of the warning device was first examined. Potential warning device positions included:

- both sides of the vehicle (visible to vehicle driver),
- same or right side of vehicle approach,
- opposite or left side of vehicle approach or
- not reported.

As seen in Figure 13 approximately 78 percent of all accidents occurred where warning devices were either on both sides of the roadway or on the same or right side of the vehicle approach. This seems counterintuitive – highway-rail crossings with warning devices on both sides of the road experienced higher accident frequencies - unless the

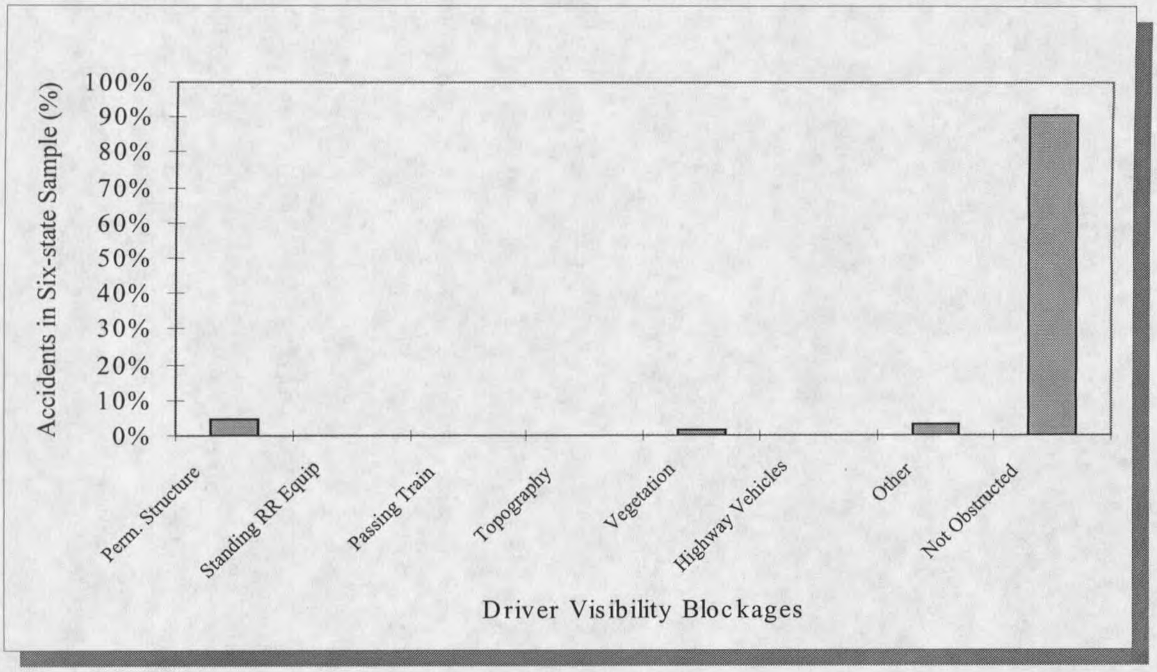

Figure 12. Highway-rail Crossing Accidents by Driver Visibility Conditions

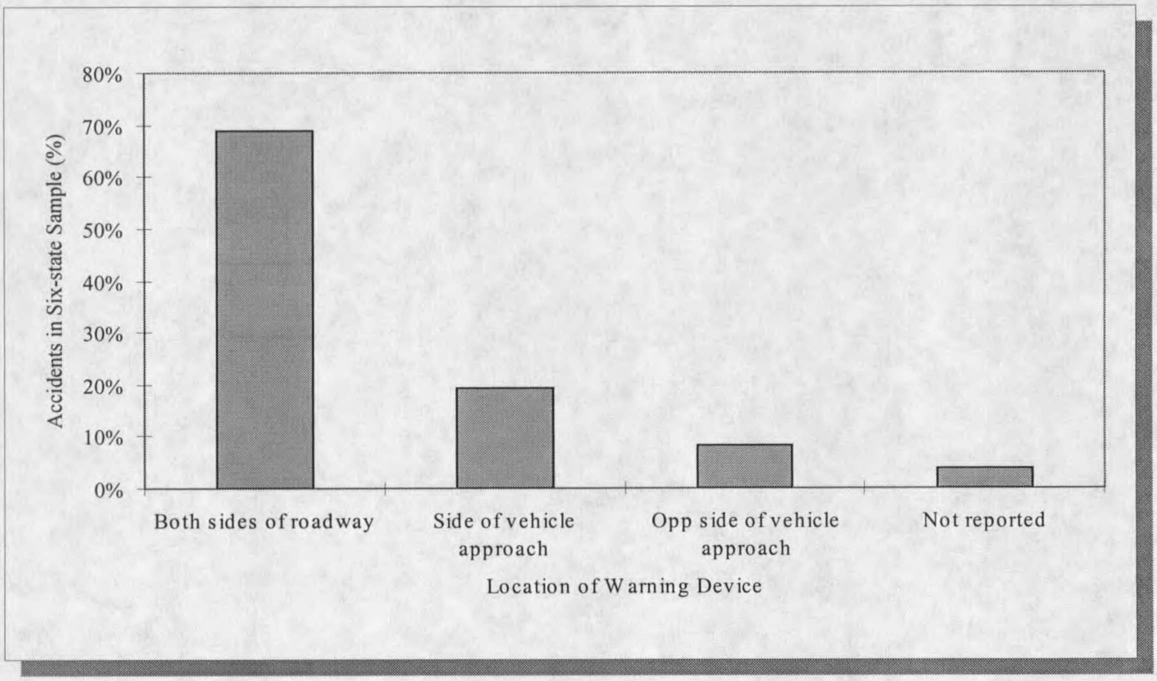

Figure 13. Highway-rail Crossing Accidents by Warning Device Location

additional warning devices were placed in response to a high number of accidents that are in fact attributable to some other cause.

At crossings where the highest level of warning device was either a gate or crossbuck, 39 percent and 37 percent of the total accidents occurred, respectively (see Figure 14). This finding was anticipated for crossbucks, which provide the lowest level of warning. The use of both gates and flashing lights would predictably result in fewer accidents, unless again placed in response to high accident frequencies attributable to other causes.

Lastly, the effect of whistle bans on highway-rail crossing safety was considered (see Figure 15). Five percent of all accidents occurred at crossings where whistle bans were in effect while 78 percent of accidents occurred at crossings with no ban in place (whistle ban conditions for the remaining 17 percent was unreported). While the five percent is not a large percentage of the total accidents, it does speak to the need to balance safety with neighborhood noise control.

Vehicle and Driver Characteristics

Vehicle and driver characteristics include vehicle and train speed at the time of the accident, motorist action just prior to the accident and driver age and gender. Most accidents occurred when vehicles were traveling at 10 miles an hour or less (68 percent) (see Figure 16). Two explanations are possible:

(1) drivers slow at a crossing but then proceed through the gates or

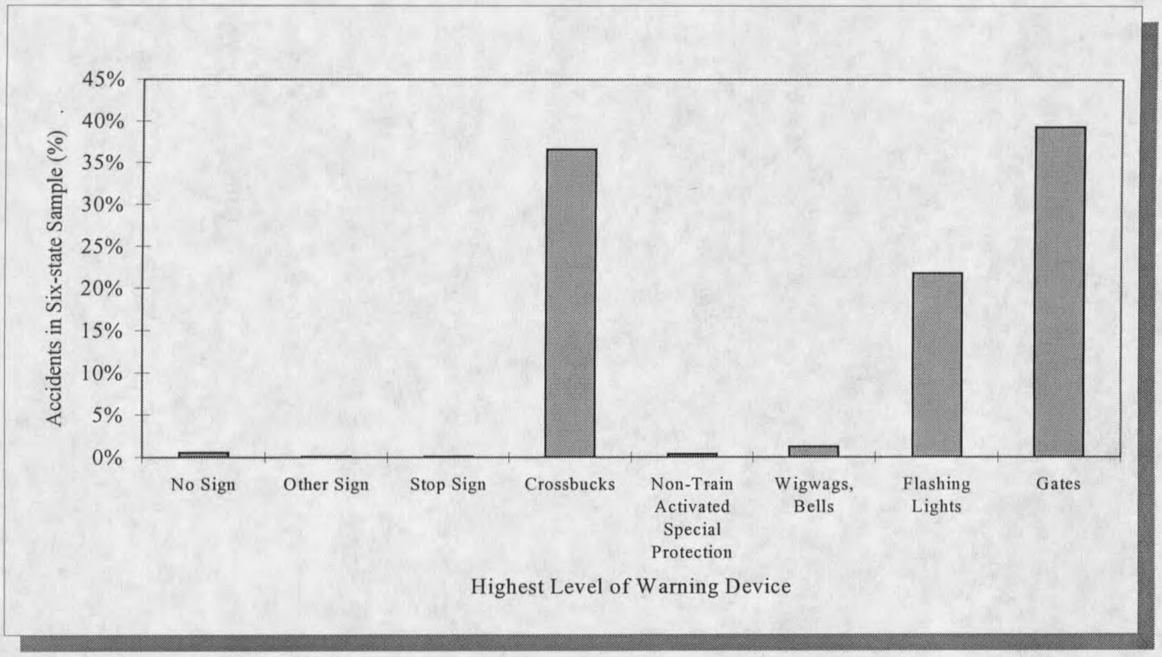

Figure 14. Highway-rail Crossing Accidents by Highest Level of Warning Device

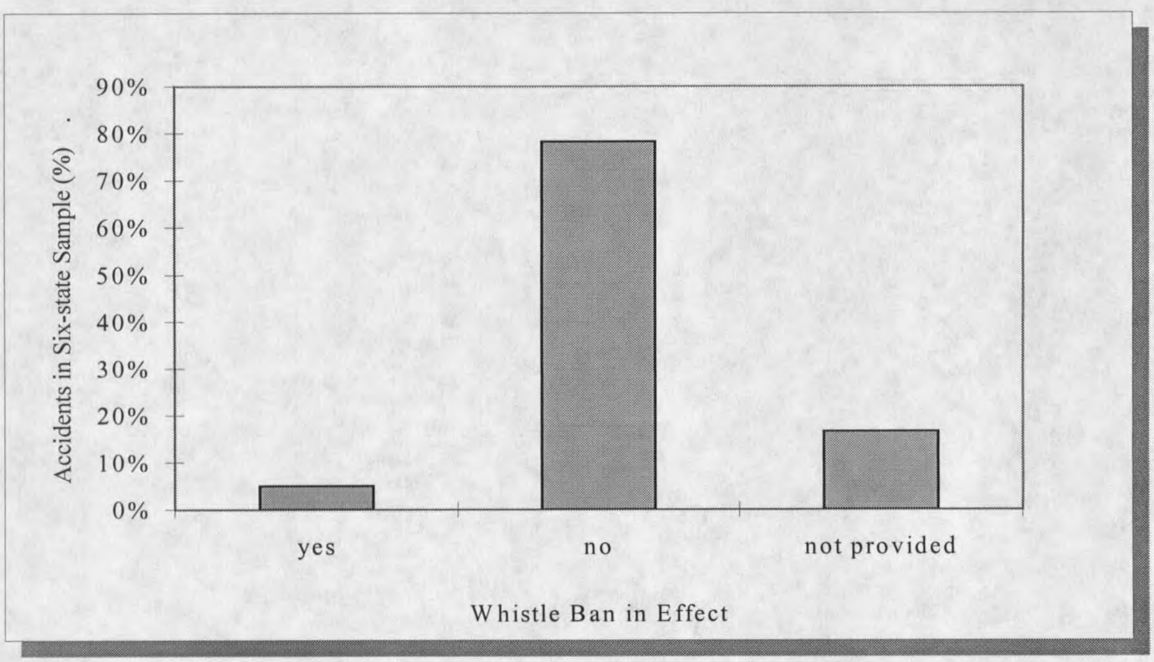

Figure 15. Highway-rail Crossing Accidents by Whistle Ban Operability

(2) drivers become conditioned to the absence of a train at a crossing and proceed without looking (i.e., negative conditioning). Negative conditioning is of particular concern in rural areas where there is little traffic and train interaction.

Not surprisingly, average train speed at the time of the accident was higher than motorist speed by an average of 17 mph to the vehicle's 11 mph. Trains, because of their heavy weight` and vehicle dynamics, aren't able to slow as quickly. It takes an average freight train traveling 55 mph with 100 rail cars a mile or more to stop (19). The higher number of accidents involving trains traveling 10 mph or less is likely reflective of the number of traffic and train interactions in urban settings. When entering a high interaction area, train operators will often slow down and exercise greater caution. This is displayed in Figure 17.

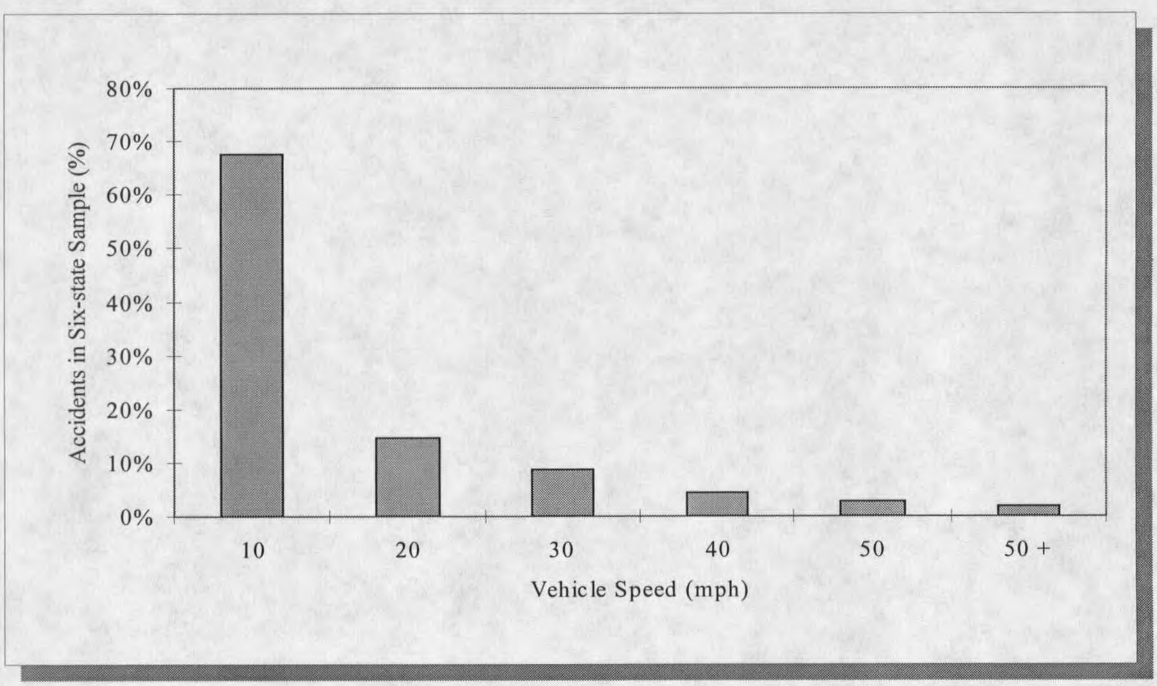

Figure 16. Highway-rail Crossing Accidents by Highway Vehicle Speed

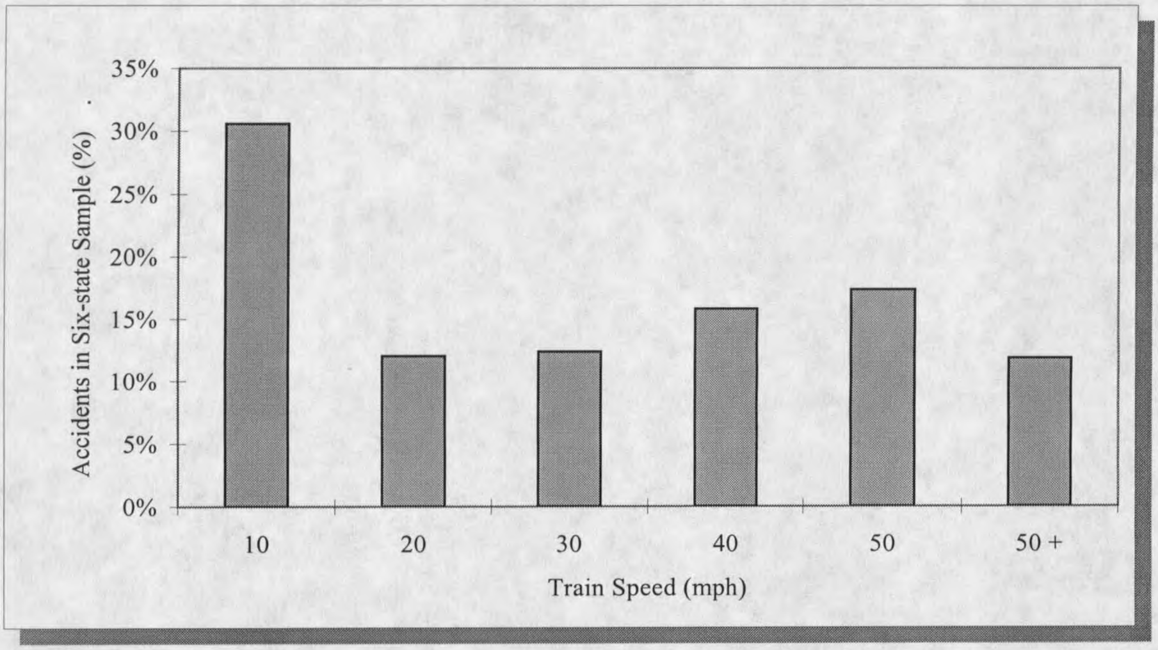

Figure 17. Highway-rail Crossing Accidents by Train Speed

Motorist actions preceding a crossing accident are categorized into the following five groups:

(1) drove around or through the gate,

(2) stopped and then proceeded,

(3) did not stop,

(4) stopped on crossing and

(5) other.

The findings related to this variable are disturbing. Nearly half of all accident-involved motorists did not stop for the crossing (49 percent) (see Figure 18).

In an attempt to better explain this behavior, driver age and gender were examined. Two peaks were noted with respect to driver age. One peak was for drivers over the age of 50, and the other for 21-30 year olds. With the "baby boomers" advancing in age, highway-rail crossing accidents involving the over-50 portion of the population could see an increase in occurrence. As seen in Figure 20, the majority of highway-rail crossing accidents also involved males (74 percent).

These statistics represent accident-specific characteristics of highway-rail crossing accidents occurring between January 1997 through December 1998 in the six-state sample. When developing the accident prediction model, much of the accident-specific detail was lost since the focus of the investigation shifted to accident frequency per crossing. Nonetheless, it is important to examine the characteristics of highway-rail crossing accidents to improve the interpretation of the causative factors.

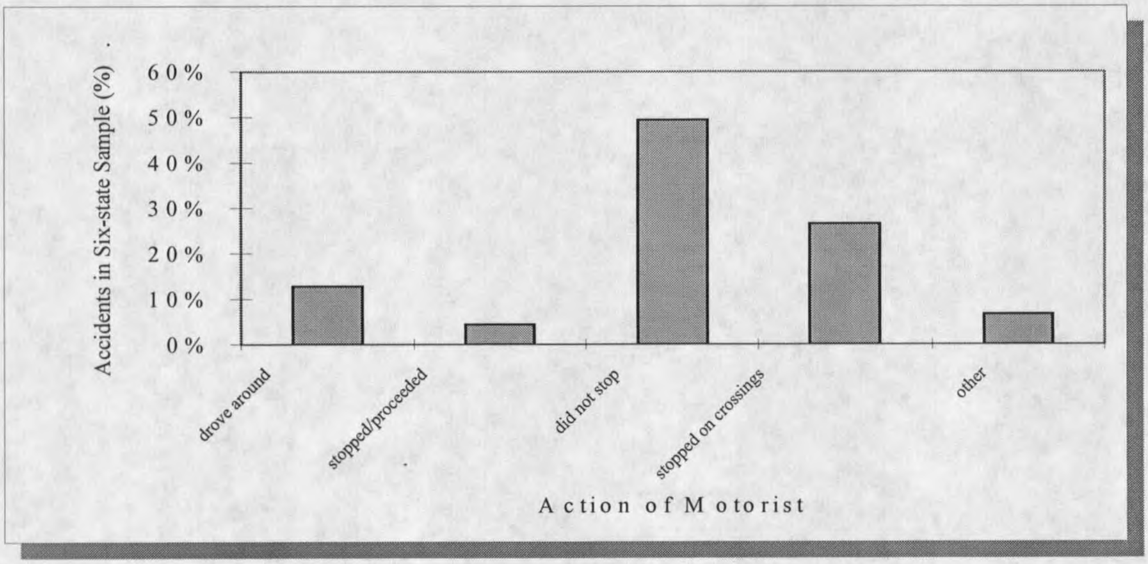

Figure 18. Highway-rail Crossing Accidents by Motorist Action

Figure 19. Highway-rail Crossing Accidents by Highway Vehicle Driver Age

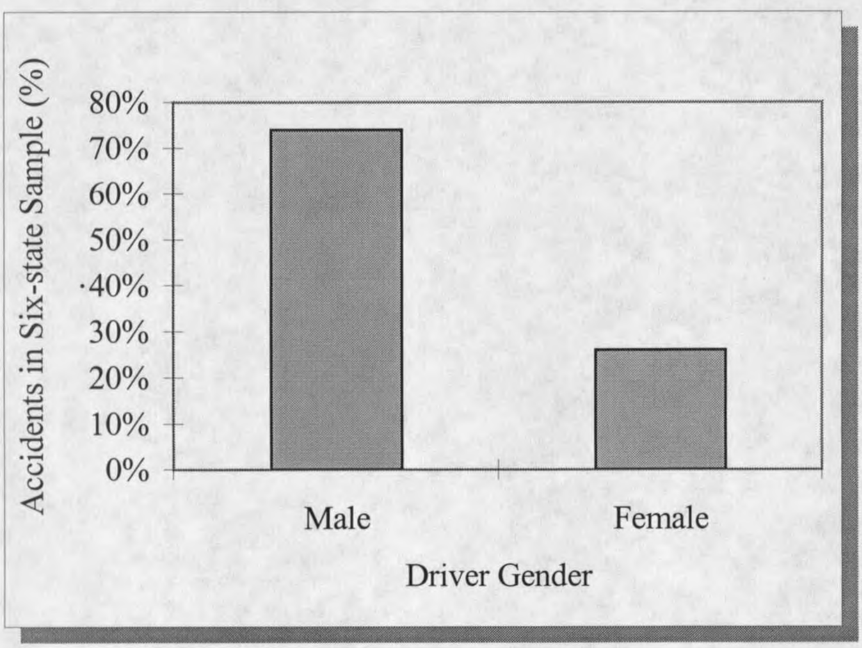

Figure 20. Highway-rail Crossing Accidents by Highway Vehicle Driver Gender

Accident Prediction Model

This section describes findings from both the Instrumental Variables correction for endogeneity among the various highway-rail crossing warning devices and accident frequency, as well as the final accident prediction model for highway-rail crossing accident frequency.

Instrumental Variables Correction for Endogeneity

As mentioned previously, the variables defining the different types of warning devices could not be used directly in the accident prediction model due to endogeneity problems with the dependent variable, accident frequency. The intent of this investigation was to determine factors that have a significant influence, either positive or negative, on the number of accidents occurring at a particular highway-rail crossing. Because warning devices likely have a significant influence on the number of accidents occurring but also may be placed in response to an already high number of accidents occurring, an inappropriate relationship results. To account for this endogeneity, the method of Instrumental Variables was used to create eight new continuous variables representing the probability of a type of warning device being in place at a particular highway-rail crossing. Binary logit models were used to determine factors significantly influencing the likelihood of warning device presence at a highway-rail crossing. The results of these eight logit models, representing each of the types of warning devices considered in this investigation, are summarized in Table 13 and detailed in Tables 14 through 21.

In Table 13, a positive sign indicates that an increase in or the presence of that variable significantly increases the likelihood of that warning device being present at the highway-rail crossing. A negative sign indicates that an increase in or the presence of that variable significantly decreases the likelihood of that warning device being present at the highway-rail crossing. The symbol 'NS' means that the variable was found to have no statistically significant influence on the presence of a particular warning device. And the symbol 'DEP' indicates the warning device variable serving as the dependent variable in the eight different models.

Tables 14 through 21 report the estimated coefficients for the various models, the standard error and the t-statistic. In each case, a t-statistic $>|1.96|$ was taken to be significant corresponding to a 95 percent confidence level.

When looking at the estimated logit models in summary (see Table 13), it is interesting to note that each of the six states sampled show a greater propensity to have crossbucks at a highway-rail crossing and a lower propensity for any other type of warning device with the exception of New York which also exhibits a greater propensity for flashing lights at highway-rail crossings. This is easily explained since crossbucks are the lowest level of warning provided and consequently, the least costly.

Table 13. Summary of Logit Model Results for Warning Devices

	MODEL 1 Crossbuck	2 Stop Sign	3 Other Sign	4 Gate	5 Flashing Lights	6 Traffic Signal	7 Wigwag	8 Bells
WARNING DEVICE								
Crossbuck	DEP	+	+	NS	-	NS	+	-
Stop Sign	+	DEP	+	NS	-	NS	NS	-
Other Sign	+	+	DEP	+	+	+	-	NS
Gate	-	NS	+	DEP	+	+	-	+
Flashing Lights	-	-	+	+	DEP	NS	-	+
Highway Traffic Signal	NS	NS	+	NS	-	DEP	-	-
Wigwag	+	-	-	-	-	NS	DEP	+
Bells	-	-	+	+	+	NS	+	DEP
CONSTANT	+	-	-	-	-	-	-	-
SPATIAL CHARACTERISTICS								
Montana	+	-	-	-	-	-	-	-
California	+	-	-	-	-	-	-	-
Texas	+	-	-	-	-	-	-	-
Illinois	+	-	-	-	-	-	-	-
Georgia	+	-	-	-	-	-	-	-
New York	+	-	-	-	+	-	-	-
TRAFFIC CHARACTERISTICS								
Number of Daily Through Trains	NS	-	NS	NS	+	+	-	-
Number of Nightly Through Trains	-	NS	+	+	-	NS	+	+
Max Timetable Speed	+	NS	-	+	+	NS	NS	+
Max Typical Speed	-	+	+	NS	-	-	-	NS
Number of Main Tracks	+	+	+	+	+	NS	+	NS
Number of Traffic Lanes	-	+	+	NS	+	+	NS	+
AADT in Both Directions	-	-	-	+	+	+	NS	+
Percentage of Trucks in Traffic	NS	NS	-	NS	+	+	+	+
ROADWAY CHARACTERISTICS								
Highway Paved or Gravel	-	+	NS	+	+	+	NS	+
75 to 150 ft to Nearest Intersection	NS	NS	NS	NS	NS	NS	NS	-
Over 200 ft to Nearest Intersection	NS	NS	NS	NS	-	NS	NS	NS
0-29 Degree Angle at Crossing	-	NS	-	-	-	+	NS	-
30-59 Degree Angle at Crossing	+	-	+	-	+	-	-	+
60-90 Degree Angle at Crossing	+	NS	NS	NS	NS	-	-	+

Table 13. Summary of Logit Model Results for Warning Devices (Continued)

	MODEL							
	1	2	3	4	5	6	7	8
	Crossbuck	Stop Sign	Other Sign	Gate	Flashing Lights	Traffic Signal	Wigwag	Bells
WARNING DEVICE								
Crossbuck	DEP	+	+	NS	-	NS	+	-
Stop Sign	+	DEP	+	NS	-	NS	NS	-
Other Sign	+	+	DEP	+	+	+	-	NS
Gate	-	NS	+	DEP	+	+	-	+
Flashing Lights	-	-	+	+	DEP	NS	-	+
Highway Traffic Signal	NS	NS	+	NS	-	DEP	-	-
Wigwag	+	-	-	-	-	NS	DEP	+
Bells	-	-	+	+	+	NS	+	DEP
ROADWAY CHARACTERISTICS								
Sight Obstruction:								
Permanent Structure	NS	NS	NS	NS	+	NS	NS	+
Topography	NS	+	NS	-	NS	NS	NS	NS
Highway Vehicle	NS	NS	NS	-	NS	NS	NS	NS
No Sight Obstructions	+	+	NS	-	NS	NS	NS	NS
Development Type:								
Open Space	+	NS	-	-	-	-	-	-
Residential	-	+	+	+	+	NS	+	+
Commercial	-	+	+	+	+	+	+	+
Industrial	-	-	+	+	NS	NS	+	+
Institutional	-	NS	+	+	+	+	NS	NS
CROSSING CHARACTERISTICS								
Surface:								
Sectional	-	+	+	+	+	-	NS	+
Full Wood Plank	-	NS	NS	-	+	NS	+	+
Asphalt	+	+	NS	NS	+	+	+	NS
Concrete Slab	NS	+	+	+	-	-	-	+
Concrete Pavement	-	+	NS	-	-	+	NS	NS
Rubber	+	-	+	+	+	+	-	+
Metal	NS	+	NS	NS	+	NS	NS	NS
Advance Warning Sign	+	-	NS	+	+	-	NS	-
Pavement Markings: Stop Lines	-	-	+	+	+	+	+	+
Railroad Crossing	+	-	NS	+	NS	-	-	+
No Markings	+	-	+	-	+	-	-	NS

Table 14. Logit Model Results for Crossbuck Warning Sign

Independent Variable	Estimated Coefficient	Standard Error	t-statistic
Constant	1.680	0.070	24.060
SPATIAL CHARACTERISTICS			
Montana	0.713	0.065	11.024
California	0.979	0.022	43.519
Texas	1.210	0.021	58.588
Illinois	0.742	0.018	42.045
Georgia	1.888	0.028	68.433
New York	1.148	0.030	37.794
TRAFFIC CHARACTERISTICS			
Number of Nightly Through Trains	-0.024	0.002	-14.717
Max Timetable Speed	0.020	0.002	12.828
Max Typical Speed	-0.018	0.002	-11.038
Number of Main Tracks	0.345	0.022	15.906
Number of Traffic Lanes	-0.086	0.013	-6.540
AADT in Both Directions	-6.56E-06	1.79E-06	-3.659
ROADWAY CHARACTERISTICS			
Highway Paved or Gravel	-0.554	0.032	-17.247
0-29 Degree Angle at Crossing	-0.081	0.041	-1.986
30-59 Degree Angle at Crossing	0.133	0.052	2.561
60-90 Degree Angle at Crossing	0.235	0.047	5.013
No Sight Obstructions	0.182	0.068	2.659
Development Type: Open Space	0.765	0.017	43.988
Development Type: Residential	-0.086	0.027	-3.158
Development Type: Commercial	-0.267	0.027	-9.958
Development Type: Industrial	-0.464	0.031	-15.177
Development Type: Institutional	-0.401	0.081	-4.960
CROSSING CHARACTERISTICS			
Surface: Sectional	-0.364	0.021	-17.634
Surface: Full Wood Plank	-0.351	0.032	-11.002
Surface: Asphalt	0.447	0.022	20.127
Surface: Concrete Pavement	-0.559	0.110	-5.079
Surface: Rubber	0.368	0.040	9.160
Stop Sign	0.682	0.047	14.375
Other Sign	0.314	0.033	9.406
Gate	-0.085	0.025	-3.340
Flashing Lights	-1.131	0.026	-43.385
Wigwag	0.734	0.080	9.121
Bells	-1.433	0.028	-51.039
Advance Warning Sign	0.372	0.025	14.823
Pavement Markings: Stop Lines	-1.050	0.018	-57.749
Railroad Crossing	0.418	0.067	6.258
No Markings	0.268	0.024	11.137
Log likelihood function	-36504.56		
Restricted log likelihood	50236.49		
Number of observations	80962		

Table 15. Logit Model Results for Stop Sign

Independent Variable	Estimated Coefficient	Standard Error	t-statistic
Constant	-3.045	0.114	-26.641
SPATIAL CHARACTERISTICS			
Montana	-2.671	0.075	-35.636
California	-2.522	0.039	-65.044
Texas	-3.339	0.043	-77.176
Illinois	-4.260	0.061	-69.676
Georgia	-1.265	0.027	-46.769
New York	-3.594	0.079	-45.323
TRAFFIC CHARACTERISTICS			
Number of Daily Through Trains	-0.013	0.002	-5.755
Max Typical Speed	0.011	0.001	10.899
Number of Main Tracks	0.448	0.036	12.495
Number of Traffic Lanes	0.055	0.025	2.169
AADT in Both Directions	-2.44E-05	4.91E-06	-4.964
ROADWAY CHARACTERISTICS			
Highway Paved or Gravel	0.172	0.046	3.730
30-59 Degree Angle at Crossing	-0.227	0.048	-4.723
Sight Obstruction: Topography	2.205	0.759	2.907
No Sight Obstructions	0.305	0.105	2.900
Development Type: Residential	0.305	0.040	7.575
Development Type: Commercial	0.207	0.044	4.725
Development Type: Industrial	-0.106	0.053	-1.987
CROSSING CHARACTERISTICS			
Surface: Sectional	0.130	0.049	2.656
Surface: Asphalt	0.727	0.036	20.048
Surface: Concrete Slab	0.449	0.104	4.294
Surface: Concrete Pavement	1.178	0.178	6.628
Surface: Rubber	-0.552	0.130	-4.239
Surface: Metal	2.355	0.573	4.108
Warning Device: Crossbuck	0.615	0.047	13.081
Other Sign	0.490	0.052	9.412
Flashing Lights	-0.711	0.053	-13.485
Wig Wag	-0.304	0.152	-2.001
Bells	-0.302	0.057	-5.266
Advance Warning Sign	-0.458	0.038	-12.217
Pavement Markings: Stop Lines	-0.106	0.032	-3.318
Railroad Xing	-0.641	0.068	-9.455
No Markings	-1.343	0.070	-19.068
Log likelihood function	-17448.26		
Restricted log likelihood	-18922.49		
Number of observations	80962		

Table 16. Logit Model Results for Other Signs

Independent Variable	Estimated Coefficient	Standard Error	t-statistic
Constant	-5.214	0.074	-70.196
SPATIAL CHARACTERISTICS			
Montana	-2.018	0.063	-31.897
California	-3.137	0.045	-69.110
Texas	-3.098	0.039	-79.203
Illinois	-1.704	0.024	-69.863
Georgia	-2.684	0.042	-63.684
New York	-2.110	0.043	-49.156
TRAFFIC CHARACTERISTICS			
Number of Nightly Through Trains	0.038	0.002	20.640
Max Timetable Speed	-0.020	0.003	-7.482
Max Typical Speed	0.032	0.003	11.885
Number of Main Tracks	0.682	0.030	22.843
Number of Traffic Lanes	0.142	0.020	7.118
AADT in Both Directions	-1.11E-05	3.25E-06	-3.410
Percentage of Trucks in Traffic	-0.006	0.002	-3.203
ROADWAY CHARACTERISTICS			
0-29 Degree Angle at Crossing	-0.151	0.075	-1.998
30-59 Degree Angle at Crossing	0.125	0.040	3.126
Development Type: Open Space	-0.412	0.029	-14.075
Development Type: Residential	0.252	0.040	6.318
Development Type: Commercial	0.412	0.040	10.394
Development Type: Industrial	0.681	0.045	14.996
Development Type: Institutional	0.449	0.119	3.785
CROSSING CHARACTERISTICS			
Surface: Sectional	0.287	0.034	8.508
Surface: Concrete Slab	0.254	0.073	3.493
Surface: Rubber	0.457	0.050	9.096
Warning Device: Crossbuck	0.342	0.033	10.364
Stop Sign	0.540	0.050	10.709
Gate	0.141	0.041	3.469
Flashing Lights	0.425	0.043	9.987
Highway Traffic Signal	0.671	0.106	6.354
Wigwag	-0.671	0.180	-3.720
Bells	0.158	0.046	3.429
Pavement Markings: Stop Lines	0.258	0.031	8.410
No Markings	0.441	0.032	13.845
Log likelihood function	-19869.74		
Restricted log likelihood	-22805.94		
Number of observations	80962		

Table 17. Logit Model Results for Gate

Independent Variable	Estimated Coefficient	Standard Error	t-statistic
Constant	-6.328	0.095	-66.779
SPATIAL CHARACTERISTICS			
Montana	-2.092	0.059	-35.656
California	-1.667	0.025	-66.858
Texas	-2.318	0.025	-92.940
Illinois	-1.458	0.020	-72.432
Georgia	-1.783	0.027	-67.115
New York	-0.387	0.028	-13.776
TRAFFIC CHARACTERISTICS			
Number of Nightly Through Trains	0.037	0.002	15.717
Max Timetable Speed	0.028	0.001	37.702
Number of Main Tracks	1.124	0.032	35.299
AADT in Both Directions	1.85E-05	2.13E-06	8.679
ROADWAY CHARACTERISTICS			
Highway Paved or Gravel	1.196	0.046	26.195
0-29 Degree Angle at Crossing	-0.218	0.044	-4.922
30-59 Degree Angle at Crossing	-0.186	0.035	-5.257
Sight Obstruction: Topography	-1.943	0.936	-2.076
Sight Obstruction: Highway Vehicle	-4.003	1.381	-2.898
No Sight Obstructions	-0.492	0.083	-5.950
Development Type: Open Space	-0.374	0.019	-19.961
Development Type: Residential	0.218	0.034	6.458
Development Type: Commercial	0.213	0.033	6.492
Development Type: Industrial	0.268	0.040	6.686
Development Type: Institutional	0.411	0.101	4.061
CROSSING CHARACTERISTICS			
Surface: Sectional	0.471	0.022	21.143
Surface: Full Wood Plank	-0.152	0.035	-4.364
Surface: Concrete Slab	0.383	0.064	5.995
Surface: Concrete Pavement	-0.700	0.181	-3.859
Surface: Rubber	0.313	0.049	6.425
Other Sign	0.171	0.042	4.099
Flashing Lights	1.135	0.031	36.556
Wigwag	-1.035	0.088	-11.805
Bells	2.323	0.030	76.357
Advance Warning Sign	0.230	0.034	6.770
Pavement Markings: Stop Lines	1.668	0.020	82.609
Railroad Crossing	0.144	0.067	2.161
No Markings	-0.612	0.069	-8.876
Log likelihood function	-23378.56		
Restricted log likelihood	-48649.86		
Number of observations	80962		

Table 18. Logit Model Results for Flashing Lights

Independent Variable	Estimated Coefficient	Standard Error	t-statistic
Constant	-4.342	0.077	-56.163
SPATIAL CHARACTERISTICS			
Montana	-1.230	0.048	-25.865
California	-1.225	0.024	-50.567
Texas	-2.109	0.024	-87.775
Illinois	-0.250	0.017	-14.510
Georgia	-1.583	0.026	-61.016
New York	0.338	0.029	11.683
TRAFFIC CHARACTERISTICS			
Number of Daily Through Trains	0.017	0.003	5.396
Number of Nightly Through Trains	-0.040	0.004	-9.486
Max Timetable Speed	0.015	0.002	7.674
Max Typical Speed	-0.005	0.002	-2.503
Number of Main Tracks	0.770	0.029	26.221
Number of Traffic Lanes	0.307	0.019	16.181
AADT in Both Directions	4.30E-05	3.12E-06	13.798
Percentage of Trucks in Traffic	0.004	0.001	2.713
ROADWAY CHARACTERISTICS			
Highway Paved or Gravel	1.110	0.038	29.461
Over 200 ft to the Nearest Intersection	-3.492	1.010	-3.458
0-29 Degree Angle at Crossing	-0.203	0.042	-4.826
30-59 Degree Angle at Crossing	0.241	0.034	7.175
Sight Obstruction: Permanent Structure	0.839	0.457	1.837
Development Type: Open Space	-0.612	0.017	-34.973
Development Type: Residential	0.177	0.031	5.801
Development Type: Commercial	0.178	0.030	6.013
Development Type: Institutional	0.623	0.100	6.233
CROSSING CHARACTERISTICS			
Surface: Sectional	0.645	0.022	29.931
Surface: Full Wood Plank	0.209	0.042	4.984
Surface: Asphalt	0.140	0.029	4.740
Surface: Concrete Slab	-0.948	0.070	-13.523
Surface: Concrete Pavement	-0.520	0.163	-3.194
Surface: Rubber	1.107	0.070	15.750
Surface: Metal	1.876	0.600	3.126
Warning Device: Crossbuck	-1.051	0.027	-39.476
Stop Sign	-0.632	0.052	-12.175
Other Sign	0.378	0.044	8.588
Gate	0.766	0.032	23.929
Highway Traffic Signal	-0.604	0.118	-5.115
Wigwag	-2.283	0.081	-28.141
Bells	2.857	0.029	98.145
Advance Warning Sign	0.516	0.030	17.146
Pavement Markings: Stop Lines	1.842	0.020	90.848
No Markings	0.228	0.029	7.808
Log likelihood function	-25175.01		
Restricted log likelihood	-54723.60		
Number of observations	80962		

Table 19. Logit Model Results for Highway Traffic Signal

Independent Variable	Estimated Coefficient	Standard Error	t-statistic
Constant	-5.884	0.351	-16.745
SPATIAL CHARACTERISTICS			
California	-3.902	0.081	-48.106
Texas	-5.642	0.142	-39.624
Illinois	-4.317	0.083	-51.765
Georgia	-5.177	0.141	-36.630
New York	-3.338	0.084	-39.883
TRAFFIC CHARACTERISTICS			
Number of Daily Through Trains	0.027	0.002	13.714
Max Typical Speed	-0.019	0.003	-7.345
Number of Traffic Lanes	0.549	0.030	18.427
AADT in Both Directions	1.51E-05	2.97E-06	5.066
Percentage of Trucks in Traffic	0.015	0.004	3.817
ROADWAY CHARACTERISTICS			
Highway Paved or Gravel	1.183	0.282	4.203
0-29 Degree Angle at Crossing	0.751	0.126	5.973
30-59 Degree Angle at Crossing	-0.629	0.176	-3.565
60-90 Degree Angle at Crossing	-0.497	0.144	-3.453
Development Type: Open Space	-3.528	0.133	-26.428
Development Type: Commercial	0.657	0.083	7.879
Development Type: Institutional	1.173	0.253	4.635
CROSSING CHARACTERISTICS			
Surface: Sectional	-0.651	0.120	-5.414
Surface: Asphalt	0.292	0.103	2.827
Surface: Concrete Slab	-0.934	0.369	-2.535
Surface: Concrete Pavement	1.837	0.229	8.018
Surface: Rubber	1.105	0.122	9.054
Other Sign	0.673	0.108	6.259
Gate	0.318	0.097	3.288
Advance Warning Sign	-0.409	0.105	-3.905
Pavement Markings: Stop Lines	0.305	0.071	4.303
Railroad Crossing	-1.236	0.117	-10.547
No Markings	-1.598	0.130	-12.247
Log likelihood function	-3207.14		
Restricted log likelihood	-4248.41		
Number of observations	80962		

Table 20. Logit Model Results for Wig Wag

Independent Variable	Estimated Coefficient	Standard Error	t-statistic
Constant	-6.931	0.291	-23.838
SPATIAL CHARACTERISTICS			
Montana	-5.549	0.356	-15.573
California	-3.406	0.079	-43.359
Texas	-5.445	0.111	-48.952
Illinois	-4.732	0.089	-53.422
Georgia	-7.663	0.412	-18.600
New York	-7.343	0.502	-14.640
TRAFFIC CHARACTERISTICS			
Number of Daily Through Trains	-0.044	0.014	-3.072
Number of Nightly Through Trains	0.039	0.015	2.538
Max Typical Speed	-0.022	0.003	-7.447
Number of Main Tracks	0.798	0.085	9.378
Percentage of Trucks in Traffic	0.030	0.003	8.863
ROADWAY CHARACTERISTICS			
30-59 Degree Angle at Crossing	-1.109	0.187	-5.927
60-90 Degree Angle at Crossing	-1.068	0.153	-6.979
Development Type: Open Space	-0.934	0.098	-9.506
Development Type: Residential	0.700	0.130	5.379
Development Type: Commercial	0.834	0.120	6.955
Development Type: Industrial	0.485	0.141	3.438
CROSSING CHARACTERISTICS			
Surface: Full Wood Plank	0.482	0.148	3.267
Surface: Asphalt	0.775	0.108	7.192
Surface: Concrete Slab	-1.711	0.513	-3.333
Surface: Rubber	-1.329	0.371	-3.578
Warning Device: Crossbuck	1.107	0.087	12.762
Other Sign	-0.807	0.194	-4.165
Gate	-1.591	0.094	-16.959
Flashing Lights	-3.050	0.087	-35.115
Highway Traffic Signal	-1.371	0.482	-2.843
Bells	6.150	0.140	43.836
Pavement Markings: Stop Lines	0.525	0.082	6.405
Railroad Crossing	-0.584	0.163	-3.593
No Markings	-1.227	0.174	-7.035
Log likelihood function	-2895.64		
Restricted log likelihood	-5007.11		
Number of observations	80962		

Table 21. Logit Model Results for Bells

Independent Variable	Estimated Coefficient	Standard Error	t-statistic
Constant	-3.792	0.101	-37.542
SPATIAL CHARACTERISTICS			
Montana	-1.468	0.051	-29.042
California	-1.034	0.023	-44.309
Texas	-1.777	0.023	-78.859
Illinois	-0.527	0.017	-30.239
Georgia	-1.794	0.026	-68.070
New York	-0.965	0.030	-32.559
TRAFFIC CHARACTERISTICS			
Number of Daily Through Trains	-0.020	0.003	-5.887
Number of Nightly Through Trains	0.060	0.005	11.986
Max Timetable Speed	0.009	0.001	10.783
Number of Traffic Lanes	0.186	0.021	8.960
AADT in Both Directions	2.36E-05	3.22E-06	7.329
Percentage of Trucks in Traffic	0.004	0.002	2.521
ROADWAY CHARACTERISTICS			
Highway Paved or Gravel	0.176	0.045	3.918
75 to 150 ft to the Nearest Intersection	-0.121	0.027	-4.417
0-29 Degree Angle at Crossing	-0.294	0.043	-6.905
30-59 Degree Angle at Crossing	0.169	0.078	2.161
60-90 Degree Angle at Crossing	0.245	0.071	3.441
Sight Obstruction: Permanent Structure	0.845	0.447	1.887
Development Type: Open Space	-0.615	0.018	-34.827
Development Type: Residential	0.299	0.038	7.884
Development Type: Commercial	0.450	0.038	11.937
Development Type: Industrial	0.291	0.043	6.724
CROSSING CHARACTERISTICS			
Surface: Sectional	0.525	0.021	24.766
Surface: Full Wood Plank	0.439	0.042	10.565
Surface: Concrete Slab	0.773	0.078	9.958
Surface: Rubber	0.307	0.057	5.343
Warning Device: Crossbuck	-1.318	0.027	-48.475
Stop Sign	-0.289	0.065	-4.471
Gate	2.215	0.031	72.180
Flashing Lights	2.988	0.030	101.062
Highway Traffic Signal	-0.370	0.132	-2.805
Wigwag	5.114	0.129	39.613
Advance Warning Sign	-0.121	0.035	-3.411
Pavement Markings: Stop Lines	1.737	0.019	91.079
Railroad Crossing	0.528	0.032	16.428
Log likelihood function	-19454.88		
Restricted log likelihood	-53555.99		
Number of observations	80962		

Other trends with respect to traffic characteristics, roadway characteristics and crossing characteristics were not as pronounced. With respect to traffic characteristics, the number of main railroad tracks, the number of traffic lanes and the percent of trucks in the traffic stream all tended to encourage the presence of many of the various warning devices. Other traffic characteristic variables showed mixed results.

Surprisingly, physical roadway and crossing features such as the distance to an intersection, the degree of angle at crossing and the presence of sight obstructions either had no significance on the presence of the various warning devices or showed mixed results. Instead, significant trends were noted related to the roadway and surface type. For paved surfaces, a greater propensity existed for nearly all types of warning devices with the exception of crossbucks. Residential, commercial, industrial and institutional development types showed a greater likelihood for most warning devices except for crossbucks that were less likely. Contrary, open space development showed a lower likelihood for most warning device presence with the exception of crossbucks that had a higher likelihood. Again, paved crossings and residential, commercial, industrial and institutional development areas likely reflect more urbanized, higher traffic areas that warrant higher levels of warning devices than crossbucks.

With a single warning device as the dependent variable, the effect of other warning devices in place at the crossing was investigated. For example, the likelihood of having a gate at a highway-rail crossing may be higher if only lower level warning devices (i.e., crossbucks) are in place at the crossing or may be lower if higher level warning devices (i.e., wig wags) were already present and gates were thought to be duplicative. Also of interest in this examination is any trends related to combinations of warning devices that are typically in place. Unfortunately, no apparent trends related to either warning device hierarchies or combinations were found. Many of the warning devices had a higher likelihood of both higher and lower level warning devices in combination.

Lastly, it is interesting to note that the number of significant variables found to influence the presence of crossbucks and flashing lights greatly exceeded the number of variables found to influence the presence of other types of warning devices. This may in fact be due to the higher frequency of use for these two types of warning devices.

Accident Prediction Model

As discussed previously in Chapter 3, prior research has shown that linear regression is not appropriate to evaluate the effect of roadway and traffic characteristics on accident frequency. Poisson or negative binomial regression provides better fitting models. Poisson regression is appropriate if the data is not overdispersed, $E[Y_i] = Var[Y_i]$. If the data is overdispersed, $E[Y_i] < Var[Y_i]$, the negative binomial model is more appropriate.

The Poisson model form was first investigated. As reported in the model output, the overdispersion parameter, α, was significant with a t-statistic equal to 6.58 implying the appropriateness of the negative binomial model form over Poisson.

The various traffic, roadway and crossing characteristics, including the probability of warning device presence, were investigated for their significance in affecting highway-rail crossing accident frequency. Again, a t-statistic $>|1.96|$ was taken to be significant corresponding to a 95 percent confidence level. Model results are provided in Table 22 and are discussed in detail below.

Traffic Characteristics. Five different traffic characteristics proved to be significant in affecting highway-rail crossing accident frequency. Higher numbers of nightly (not total) through trains and the average annual daily traffic (AADT) in both directions were both found to increase highway-rail crossing accident frequency. This is directly intuitive as higher train and traffic volumes lead to higher potentials for conflict at crossing points.

Table 22. Accident Prediction Negative Binomial Results

Independent Variable	Estimated Coefficient	Standard Error	t-statistic
Constant	-6.719	0.136	-49.498
TRAFFIC CHARACTERISTICS			
Number of Nightly Through Trains	0.039	0.005	8.236
Max Timetable Speed	0.021	0.002	12.828
Number of Main Tracks	0.484	0.064	7.556
Number of Traffic Lanes	0.170	0.031	5.418
AADT in Both Directions	3.59E-05	3.77E-06	9.524
ROADWAY CHARACTERISTICS			
Highway Paved or Gravel	0.295	0.090	3.259
CROSSING CHARACTERISTICS			
Surface: Sectional	0.260	0.071	3.684
Surface: Full Wood Plank	0.312	0.074	4.233
Pavement Markings: Stop Lines	0.747	0.073	10.196
Probability of a Stop Sign	19.615	2.174	9.024
Probability of a Gate	-2.974	0.202	-14.687
Probability of Flashing Lights	1.075	0.182	5.922
Probability of a Highway Traffic Signal	-114.447	23.651	-4.839
Probability of Bells	0.649	0.170	3.820
Log likelihood function	-7127.55		
Restricted log likelihood	-7166.86		
Number of observations	80962		

Also, the greater the number of main track lines and traffic lanes at the crossing, the higher resulting accident frequency. This finding is most likely related to the previous; higher train and traffic volumes require a greater number of tracks and traffic lanes to operate.

Lastly, the higher the defined maximum timetable train speed, the higher the predicted accident frequency. As previously discussed, trains require extensive stopping distances. At higher speeds, these stopping distances extend. Trains traveling at lower speeds may be able to see an obstruction ahead and slow sufficiently to prevent an accident whereas trains traveling at higher speeds may not.

Roadway Characteristics. Only one variable in this category proved significant in predicting accident frequency. If a highway is paved, there is a higher likelihood of an accident than if it is gravel. Again, this may be a reflection of earlier findings; paved roads most often exist in higher density areas that experience higher train and traffic volumes. Hence, the accident exposure is increased.

Surprisingly, development type, roadway geometry (i.e., vicinity intersections, approach angles, etc.) and sight obstructions were not found to be significant factors in affecting highway-rail crossing accident frequency. This speaks to the design-related issues of this investigation. Design improvement recommendations cannot be made given the above findings since no design-related factors were found to either positively or negatively affect accident frequency at highway-rail crossings.

<u>Crossing Characteristics</u>. Rather than focusing on design-related improvements, when looking at the significant crossing characteristics, one may want to consider improvements in the use of warning devices. The presence of stop signs, flashing lights or bells all increase the predicted accident frequency. The presence of gates and highway traffic signals significantly reduces the accident frequency.

These latter findings are more easily explained. Gates provide a physical blockage that serves as a deterrent to crossing. Highway traffic signals are most commonly present in higher density, higher traffic areas that may physically prevent illegal movement and carry penalties for violation of the signal indication. The commonality of stop signs and the subsequent desensitization of motorists to the sign requirements may explain the unwanted effect of stop signs at highway-rail crossings. Similar reasoning supports the finding that stop lines at a highway-rail crossing resulted in higher accident frequencies. The likelihood for higher accident frequencies resulting from the presence of active warning devices such as flashing lights and/or bells is not as expected.

Two different crossing surfaces, sectional and full wood plank, were also found to increase the frequency of accidents at highway-rail crossings. The reasons for this are not immediately intuitive. One would suspect that these surface types are most frequently installed at low-volume highway-rail crossings due to their higher maintenance requirements. Lower volume crossings would predictably lead to lower accident exposure but this was not the case.

CHAPTER 5

CONCLUSIONS AND RECOMMENDATIONS

Recall that the intent of this investigation was to respond to the three-part problem detailed earlier in this document.

(1) Highway-rail crossing accidents continue to result in a high number of fatalities annually and therefore require further investigation beyond the state-of-the-practice.

(2) A high proportion of highway-rail crossing accidents occur at locations where active warning appurtenances are in place suggesting that existing strategies for improving safety are ineffective and require re-examination.

(3) Existing safety evaluation methods (i.e., accident prediction models) do not adequately describe design or other characteristics that are most detrimental to highway-rail crossing safety thus preventing the prioritization or targeting of safety improvements.

Despite an encouraging trend toward improved safety, highway-rail crossings still result in a significant number of fatalities per year with a high associated economic cost related to lost productivity and costs associated with property damage, medical care, insurance, funeral requirements, legal activities and other. Perhaps most disturbing is that over 50 percent of highway-rail crossing accidents occurred at public crossings where active warning devices (i.e., gates, lights, bells, etc.) were in place and functioning properly (4).

With respect to safety evaluation methods, the U.S. Department of Transportation's (USDOT) Accident Prediction Formula, developed in the early 1980's, is most widely used although three other predominant accident prediction models exist: the Peabody Dimmick Formula, the New Hampshire Index and the National Cooperative Highway Research Program (NCHRP) Hazard Index.

The USDOT Accident Prediction Formula is a complex, three-part formula that comprehensively addresses characteristics that may influence a crossing's level of safety (i.e., train and traffic volumes, site and surface characteristics, road/rail-side appurtenances, etc.). However, the formula does not readily provide the magnitude to which each of the characteristics contribute to a crossing's level of safety. This shortcoming makes it difficult to identify or prioritize design or improvement activities that will most effectively address safety-related problems. Further, a steady reduction over time is occurring in the normalizing coefficients used in the formula implying a steady decline in accident prediction model accuracy as compared to observed values.

The Peabody Dimmick Formula, also referred to as the Bureau of Public Roads Formula, was developed in 1941 and is used to predict the number of accidents over a five-year time period. When it was developed, the Peabody Dimmick Formula was based on accident data from rural crossings in 29 states. Non-representative sampling (only rural crossings) hinders the equation's validity. The age of the formula also presents a problem with respect to its ability to predict accidents at crossings where more recent technology is being used.

The New Hampshire Index is a simple, multiplicative relationship between annual average daily traffic, average number of trains per day and a protection factor indicative of the warning devices present at a highway-rail crossing. This basic formula has been modified significantly to form several different variations used among various states. The dissimilarity between the New Hampshire Index model variations raises concerns over its validity. While most of the discrepancies can be attributed to state preferences, concern is raised due to the lack of consistency. Depending on the variation chosen, prediction values vary considerably.

The National Cooperative Highway Research Program (NCHRP) Hazard Index, documented in NCHRP Report 50, was published in 1964 in a joint effort between the American Association of State Highway Officials (AASHO now AASHTO) and the Association of American Railroads (AAR) in response to the disproportionately high number of accidents occurring at highway-rail crossings. The NCHRP Hazard Index used accident data that spanned five years and was collected by the Interstate Commerce Commission, state agencies and others (9). The NCHRP Hazard Index is concise and easy to use. There are only three variables to calculate which provides this ease of use, but this limits it's descriptive capabilities. In addition, different factor values exist depending on whether the highway-rail crossing is classified as urban or rural although little guidance exists to accurately define those classifications.

Given the shortcomings of the previously identified accident prediction models, this investigation sought to develop an improved accident prediction model that was simple to use and interpret, yet descriptive.

Data to support this investigation came from two sources: (1) the Federal Railroad Administration's Office of Safety accident/incident database and (2) the Federal Railroad Administration's Office of Safety highway-rail crossing inventory. Because of the high number of crossings in the U.S., the data for this investigation was limited to a six-state sample selected by geographical location, number of crossings, and the variability associated with crossings and a two-year time period from January 1997 to December 1998. The states selected for inclusion were California, Montana, Texas, Illinois, Georgia, and New York. Only public highway-rail crossings were considered in this investigation.

The accident prediction model developed as part of this investigation utilizes the combined FRA databases related to accidents/incidents and highway-rail crossing inventory information. The intent was to develop a model capable of predicting the frequency of highway-rail crossing accidents on the basis of various site and traffic (highway and rail) conditions.

Prior research has shown that linear regression is not appropriate to evaluate the effect of roadway and traffic characteristics on accident frequency. Poisson or negative binomial regression provides better fitting models. Poisson regression is appropriate if the data is not overdispersed, $E[Y_i] = Var[Y_i]$. If the data is overdispersed, $E[Y_i] < Var[Y_i]$, the negative binomial model is more appropriate.

The appropriateness of the negative binomial model form was confirmed; the overdispersion parameter, α, was significant with a t-statistic equal to 6.58. The various traffic, roadway and crossing characteristics, including the probability of warning device presence, were investigated for their significance in affecting highway-rail crossing accident frequency. A t-statistic $>|1.96|$ was taken to be significant corresponding to a 95 percent confidence level.

In looking at the model results, many of the findings were as anticipated. Higher numbers of nightly (not total) through trains and the average annual daily traffic (AADT) in both directions were both found to increase highway-rail crossing accident frequency as were the number of main track lines and traffic lanes at the crossing. Higher train and traffic volumes require a greater number of tracks and traffic lanes to operate and lead to higher potentials for conflict at crossing points. Higher maximum timetable train speeds also resulted in higher predicted accident frequencies likely attributable to train stopping distances.

When considering roadway characteristics, only one variable proved to be significant. If a highway is paved, there is a higher likelihood of an accident than if it is gravel. Surprisingly, development type, roadway geometry (i.e., vicinity intersections, approach angles, etc.) and sight obstructions were not found to be significant factors in affecting highway-rail crossing accident frequency. This is a significant finding for the design-related issues of this investigation. Design improvement recommendations cannot be made given the above findings since no design-related factors were found to either increase or decrease accident frequency at highway-rail crossings.

Rather than focusing on design-related improvements, one may want to consider improvements in the use of warning devices. The presence of stop signs, flashing lights or bells all increase the predicted accident frequency. The presence of gates and highway traffic signals significantly reduces the accident frequency.

These latter findings suggest that the most effective highway-rail crossing warning/traffic devices are those that include some form of physical blockage or carry with them a strong and frequently enforced penalty for violation. Gates provide a physical blockage that serves as a deterrent to crossing. Highway traffic signals are most commonly present in higher density, higher traffic areas that may physically prevent illegal movement and carry penalties for violation of the signal indication.

The commonality of stop signs and the subsequent desensitization of motorists to the sign requirements may explain the unwanted effect of stop signs at highway-rail crossings. Similar reasoning supports the finding that stop lines at a highway-rail crossing resulted in higher accident frequencies. The likelihood for higher accident frequencies resulting from the presence of active warning devices such as flashing lights and/or bells is not as expected.

Table 23 compares these findings with those of the previously developed accident prediction models. As discussed previously, the Peabody Dimmick Formula, the New Hampshire Index (in its original form) and the NCHRP model are all simple to apply but lack descriptive capabilities due to limited factor considerations. The similarities among these three models is readily apparent in Table 23. Surprising to note is the similarities between the USDOT Accident Prediction Model and the Negative Binomial model developed as part of this investigation with respect to the factors influencing highway-rail crossing accident frequency. Both found the number of trains and train speeds to be significant factors in determine highway-rail crossing accident frequency though slight variations exist with respect to the specific factors considered (i.e., the USDOT Accident Prediction Model considers the number of day through trains and the total number of trains while the number of nightly through trains was found to be a significant factor using the negative binomial model form). Also, both the USDOT Accident Prediction Model and the negative binomial model developed here found highway surface – paved or gravel – to be a significant factor affecting accident frequency. The presence of various warning devices at a crossing were also found to be significant in both models though the effect of various warning devices on accident frequency was sometimes in opposition between the two models. For example, the presence of flashing lights at a highway-rail crossing is assumed to reduce accident frequency in the USDOT Accident Prediction Model whereas they were shown to have the opposite effect in the negative binomial model. Further, the negative binomial model found crossing surface type to significantly affect the frequency of accidents at highway-rail crossings.

The similarity between the USDOT Accident Prediction Model and the negative binomial model developed here suggests that in fact a successful alternate model has resulted capable of predicting accident frequencies at highway-rail crossings. The benefit to be gained through the development of this alternate model is: (1) a greatly simplified, two-step estimation process, (2) comparable supporting data requirements, and (3) interpretation of both the magnitude and direction of the effect of the factors found to