DEVELOPMENT OF MODELS TO QUANTIFY
THE IMPACTS OF SIGNALI ZATION
ON INTERSECTION CRASHES

Prepared for:

Florida Department of Transportation

January 2002

DEVELOPMENT OF MODELS TO QUANTIFY THE IMPACTS OF SIGNALI ZATION ON INTERSECTION CRASHES

Written by

Juan C. Pernia
Research Associate

Jian John Lu, Ph.D, P.E.,
Associate Professor

Michael X. Weng, Ph.D.
Associate Professor
Department of Industrial and Management Systems

Xiangli Xie,
Research Assistant

and

Zhuo Yu
Research Assistant

Transportation Program
Department of Civil and Environmental Engineering
University of South Florida,
4202 E. Fowler Ave., ENB 118,
Tampa, Florida 33620
Phone: (813) 974-5817
Fax (813) 974-2957.

Presented to
Florida Department of Transportation

A Project Sponsored by
Florida Department of Transportation

January 2002

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Traffic signals have been considered a way to improve traffic safety and operations at intersections where the warrants for traffic signal installation, specified by the Manual on Uniform Traffic Control Devices (MUTCD), are met. However, the impacts of signalization on crashes at intersections are complicated and have not been investigated in depth. This research focused in the evaluation of the impacts of signalization on crashes at newly signalized intersections in Florida through the development of statistical crash prediction models that can estimate the expected number of crashes at an intersection before and after the installation of traffic signals, in terms of total number of crashes and number of crashes for different crash types, including angle, left-turn, rear-end and other crashes. In the research, a before and after analysis was also performed for number of all crashes, different type of crashes, different crash severities, and crash rates, as well as an evaluation of the impacts of signalization on crashes through a case based crash prediction system. The original crash database used in this research was taken from the Florida crash database maintained by FDOT. It consists of all crashes occurred on state roads within a ten-year period from 1989 to 1998. This database is updated yearly and includes all long form reported crashes with a fatality, an injury, and high property damage occurred on state roads.

The first part of the project focuses on a before-and-after analysis to compare the number of crashes and crash rates based on different crash types, crash severities and surrounding land uses. Distribution fitting for Poisson distribution or Negative Binomial distribution was performed based on crash data. From the distribution fitting, the 50^{th} and 85^{th} percentile values were estimated and compared between the before and after period. The annual average number of crashes and crash rates were also compared to explore the safety impact of signalization on intersection crashes. Paired t-test was employed to determine if there was a statistically significant difference between both periods.

On the second part of the research, statistical crash predictive models were developed to estimate the average number of crashes as well as the corresponding variances in terms of all crashes and specific type of crashes at intersections before and after the installation of

traffic signals. During the modeling process, Poisson regression was first conducted as the initial step for each model, with negative binomial regression being applied where the crash data showed over-dispersion. The regression parameters were estimated by using maximum likelihood method with Statistical Analysis Software. The goodness-of-fit of developed models were evaluated based on Pearson's R-square and likelihood ratio index.

In the third part, an operational research approach denominated case-based crash prediction system was used to predict crash frequencies at new intersections based on some known cases. In this method, the most similar intersections with respect to roadway environment for application to a new intersection were retrieved from a training database. Then, the information and knowledge from the previous cases were adapted or reused to solve the new case, and the predicted crash frequency for the new intersection was evaluated. Once this system was ready, a testing database was used to estimate the number of crashes for intersections with specific characteristics. Lognormal modeling was performed to obtained the final results for this new approach.

With the models developed during the research, the average number of crashes at an intersection before and after the installation of a traffic signal can be estimated given the intersection characteristics. The change (increase or decrease) of the estimated crash frequencies before and after signalization can be calculated, using either the tables found in the appendices or the developed models, to represent the impacts of signalization. Based on the results of the crash data analysis (before-and-after comparison of mean values), it was concluded that signalization did have some impacts on traffic safety at intersections. All the following results were statistically significant at a 95% confidence level.

- Non-injury number of crashes and non-injury crash rates would increase, fatal crash rates would decrease, and number of injury crashes would increase after signalization.
- Total number of crashes and total crash rates, as well as number of rear-end crashes and rear-end crash rates, would increase after signals were installed, while

number of angle and left turn crashes, angle and left turn crash rates, and right turn crash rates would be reduced.

- Number of all crashes and all crash rates would be increased in urban areas, and crash rates for rural areas would be reduced.

For statistical models and operational research models, different variables of the intersections would be related to the occurrence of crashes depending on the crash types considered. The estimated impacts on crashes of different characteristics of intersections could be found on the result tables on the appendices.

ACKNOWLEDGMENT

This project was sponsored by the Florida Department of Transportation. This support is greatly appreciated. The authors would like to express special thanks to Mrs. Deborah Snyder, District Traffic Operations Engineer in District One of the Florida Department of Transportation, for her technical support and guidance.

Thanks are given to Yue Shen, graduate research assistant of the Department of Industrial and System Management Engineering, for his contribution on phase three of the project.

CHAPTER 1. INTRODUCTION

1.1. Background

Traffic crashes are an inevitable, however undesirable, transportation outcome. They cause the loss of wages, time, productivity, and especially loss of human lives, for which value cannot be estimated. Each year, hundreds of thousands of traffic crashes occurred in the United States. As an example, in 1999 there were 6,279,000 estimated traffic crashes according to the Traffic Safety Facts 1999. Out of this estimated number of crashes, 37,043 were fatal crashes with 41,611 fatalities. For the state of Florida, specifically, a total of 2,626 fatal crashes caused 2,918 fatalities in 1999. These figures clearly illustrate why traffic safety problems are a major concern to public, politicians and transportation professionals. These figures also indicate that there is a tremendous need for improving traffic safety, especially at intersections, where crashes happen more frequently as compared to roadway segments.

Intersection-related crashes make up a very high percentage of the total number of crashes in the roadway system. For example, Figure 1.1 presents the national statistics for crashes by location and crash severity for 1999. For all fatal crashes, 22.98% occurred at intersections or intersection-related locations. In regard to all traffic crashes, 44.69% occurred at intersections or intersection-related locations. For injury crashes, the percentage is close to 50%, and for property damaged only (PDO) crashes over 42%. The main reason for these high percentages is that intersections are areas shared by two or more roads, where roadway users including vehicle drivers, cyclists, and pedestrians have to make a decision or are confronted with many choices to make, whether to stop or keep going, go left, right or straight, etc. The complexity of movements of vehicles at intersections results in the basic problem for intersections, too many conflict points. Usually, once a traffic conflict is not avoided, a traffic crash will occur. Therefore, safety analyses at intersections are necessary.

A valid approach to address safety at intersection is through intersection crash studies. An appropriate intersection crash study is to explore the different crash patterns before and after signalization based on intersection data. This research was conducted by using this approach to find out the crash patterns before and after the signalization, and then based on the different patterns, the impact of signalization on intersection crashes was found out. Furthermore, crash prediction models were developed to quantify the impacts of signalization on intersection crashes.

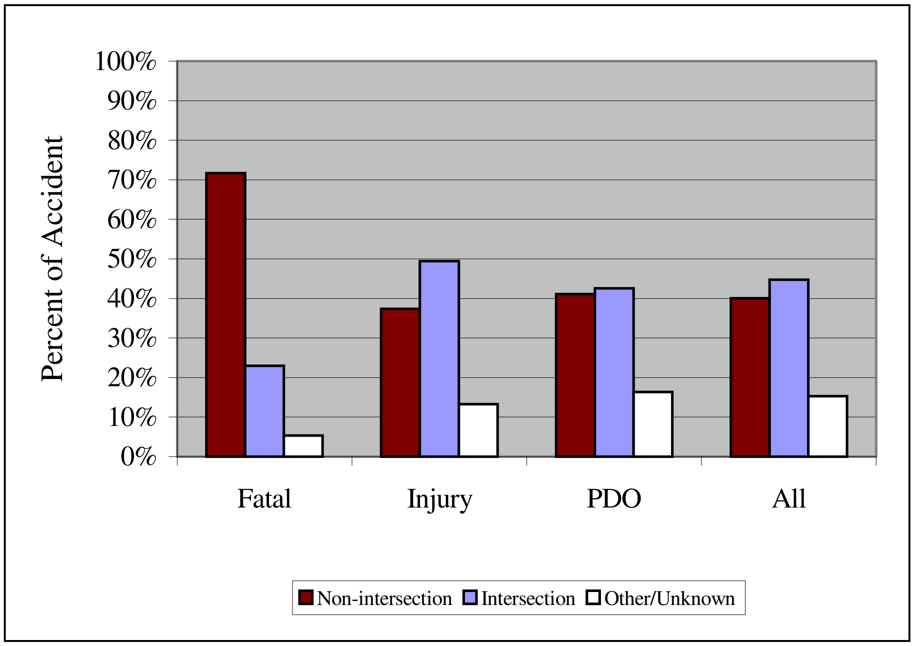

Figure 1.1. National Statistics of Accidents by Location and Severity
(Source: Traffic Safety Facts 1999)

It is very important to develop crash prediction models to estimate intersection crash frequencies. However, crash frequency prediction is not an easy task due to the large number of factors that affect crash occurrences and possible complicated interactions among them. These factors can be grouped into five categories: drivers, traffic, intersection or roadway segment, vehicles, and environment (e.g. weather condition). Although four of the five factors play an important role in traffic safety, traffic engineers can only directly manage factors related to roadway through intersection design or improvement phases.

Within roadway factors, traffic controls at intersections are very important. These traffic controls include yield sign, stop sign, flashing beacon and traffic signals. Traffic signals are in the higher level of these controls. Furthermore, it has long been thought by the general public that traffic signals could reduce the number and/or severity of crashes at an intersection. However, many traffic engineers and traffic operations professionals know this is not necessarily the case. Results from many studies showed that the number of crashes could increase and crash rates may not have a significant decrease after signal installation. Moreover, previous research also indicates that rear end crashes will increase after the installation of a traffic signal while angle crashes will decrease. Rear end crashes will increase after a signal is installed because more vehicles will have to stop on the major road than before. Angle crashes will decrease because the traffic signal will give the vehicles from the minor road the right of way when crossing the major road. However, the effects of signal installation on traffic safety at intersections have not been fully investigated, especially in Florida. Therefore, this research focused in the evaluation of the impacts of signalization on crashes at newly signalized intersections in Florida through the development of statistical crash prediction models that can estimate the expected number of crashes at an intersection before and after the installation of traffic signals, in terms of total number of crashes and number of crashes for different crash types, including angle, left-turn, rear-end and other crashes. A before and after analysis for number of all crashes, crash types, crash severities, and crashes by surrounding land use, as well as for crash rates, and an evaluation of the impacts of signalization on crashes through a case based crash prediction system were also performed. In the before-and-after comparison, statistical tests were performed to evaluate the significance of the increase or decrease of crashes or crash rates.

Furthermore, it had long been thought that a crash prediction model to predict the number of crashes at an intersection before and after signalization was sorely needed. The models of this research could be used by traffic operations and safety engineers to estimate the changes in total number of crashes and number of crashes by type when they are faced with the decision whether or not a traffic signal is in the best interest of the public. These models would also be extremely helpful to inform politicians and public what could happen once a traffic signal is installed. As an example, the possible impacts of a signal on crashes at an intersection could be used as additional information during the analysis of signal warrants at an intersection when considering signalization as an alternative solution for the intersection's problems. These impacts, which could be estimated with the models by predicting the number of crashes before and after signalization, may illustrated the increase or decrease of specific type of crashes depending on the conditions of the intersection.

This research project not only developed crash prediction models but also covered other related aspects. The project was divided into three phases: before and after comparison study, development of statistical crash prediction models, and development of case-based crash prediction models. In the first phase, the following tasks were completed: 1) collecting the specific crash data from the original FDOT crash database, and creating a crash database for the intersections considered in the research; 2) conducting a before-and-after comparison analysis to determine the change in the number of crashes and crash rates after traffic signals were installed at intersections; and 3) Poisson and Negative Binomial distributions were used to fit the observed crash distributions, and the 50th percentile and 85th percentile values were gathered from the fitted distributions matching with the observed data. Because the crash counts are discrete numbers, Poisson distribution was usually used to analyze the number of crashes and crash rates to investigate the impact of signalization on intersection crashes for this phase. Negative Binomial distribution was also used due to the fact that recent studies indicated that this distribution would be more accurate to fit the crash distribution when the data is over dispersed [Nicholson (1985), Poch et. al. (1996)].

In the second phase of the project, the preliminary database was further reduced and processed to generate the final database for statistical modeling purposes. The modeling procedure can be summarized into five steps: (1) finding an appropriate probability function to describe the random variation of crash frequencies; (2) determining an appropriate functional form and parameterization to describe the effects of independent variables on the expected crash frequencies; (3) selecting the right independent variables to include and collect associated data; (4) estimating the regression parameters using appropriate statistical algorithm based on the crash data and the probability assumptions; and (5) assessing the quality of the model to make sure that the developed model makes good engineering sense as well as fulfilling corresponding statistical goodness-of-fit criteria. For modeling, five types of crashes were selected: (1) number of all crashes per year; (2) number of rear-end crashes per year; (3) number of angle crashes per year; (4) number of left-turn crashes per year; and (5) number of other crashes per year (including all other crash types). The reason to combine other types of crashes together was the insufficient crash counts for each crash type to perform separate modeling analysis. For each of the cases considered, two models were developed, one based on data before signalization, and the other based on data after signalization. In the model developing process, the Poisson regression was used as an initial step, with the negative binomial model then being applied where over dispersion existed in the crash data. In regard to predictor variables for the model, a total of seven characteristics related to intersections were selected, including the average daily traffic (ADT) of the major road, urban/rural, land use of surrounding area, number of lanes on major road, posted speed on major road, type of median, and shoulder treatment. The maximum likelihood method was used to estimate the regression coefficients. The methods applied to test the goodness-of-fit of the models include Pearson's residual, Pearson's R-square, and likelihood ratio index. Once the models were developed, the expected number of crashes at an intersection before and after the signalization were estimated by using the "before" model and the "after” model. Then, the changes of the estimated number of crashes were estimated as impacts of signalization. These estimated results were finally tabulated in order to present a simple and clear overview of the impact of signalization on crashes for intersections with different characteristics.

The third phase of the project consisted of a new approach (operational research) to predict crash frequencies at a new intersection based on some known cases. The basic idea of this new approach is to remember old solutions (crash frequencies) to similar problems (intersections) and to adapt them to fit a new problem (intersection) rather than having to solve it from scratch. This method, denominated case-based crash prediction system, involves the following basic steps: (1) retrieving from a training database the most similar known intersections with respect to roadway environment for application to the new intersection, (2) adapting or reusing the information and knowledge from the previous cases to solve the new case, (3) evaluating the proposed solution (crash frequency) to the new case. This case-based crash prediction system was evaluated to know what was going right and what was going wrong. Once this system was ready, a testing database was used to estimate the number of crashes for intersections with specific characteristics.

1.2. Crashes at Intersections

In reference to crashes, there are many different types that can occur at an intersection. One of the most common types is the rear end crash, which usually occurs when one vehicle collides with another vehicle in the "rear end" of the vehicle. Angle crashes are also common at intersections where one vehicle tries to cross the path perpendicular to the other vehicle. Left turn and right turn are similar to angle crashes except that one vehicle is making a turn of some sort when they cross the path of the other vehicle. Sideswipe is another type of crash that can occur at an intersection, which usually happen when one vehicle attempts to change lanes and collides with another vehicle on the side. Finally, crashes related to pedestrians are also important to be considered at the intersection. There are some other types of crashes that can occur at an intersection but these crash types rarely happen. Table 1.1 presents all the different types of crashes.

In reference to these types of crashes at intersections, several previous studies concluded that rear end crashes would increase significantly after signal installation, and angle crashes would decrease significantly after signal installation [King et. al (1975), Short et. al. (1982), Shen (1984), Radwin et. al. (1987)]. As for left turn crashes, different researches have different results in regard to the change of number of crashes and crash rates after the traffic control is changed [King et. al (1975), Radwin et. al. (1987)]. This research investigated the impacts of signalization in total number of crashes and in several types of crashes based on data collected only in Florida. Therefore, it can give a very important insight on crashes at intersections.

1.3. Modeling

1.3.1. Statistical Properties of Intersection Crash Frequencies

Traffic crashes are random and discrete events that are sporadic in nature, and obviously, crash frequencies and crash rates are necessarily non-negative. In fact, crash frequencies for particular intersections or relatively short roadway segments at a time interval are typically small integers. Furthermore, it is not uncommon for a substantial proportion of locations in a crash study to have no crashes at all during the study period. Also, crash frequency data show great variation. These variations are clearly consistent with the complex traffic crash mechanics, which includes pure randomness and the interactions of five major factors: drivers, traffic, intersection or roadway segment, vehicles and environment.

Table 1.1. Crash Types

Crash Code Number	Crash Type
1	Rear End
2	Head On
3	Angle
4	Left Turn
5	Right Turn
6	Sideswipe
7	Backed Into
8	Parked Car
9	W/Other Motor Vehicle on Road
10	Pedestrian
11	Bike
12	Bike in Bike Lane
13	Moped
14	Train
15	Animal
16	Sign/Sign Post
17	Utility/Light Pole
18	Guardrail
19	Fence
20	Concrete Barrier Wall
21	Bridge Abutment/Pier
22	Tree/Shrub
23	Construction Barricade/Sign
24	Traffic Gate
25	Crash Attenuators
26	Fixed Object Above Road
27	Other Fixed Object
28	Moveable Object on Road
29	Ran Into Ditch/Culvert
30	Ran Off Road Into Water
31	Overturned
32	Occupant Fell From Vehicle
33	Tractor Trailer Jack-knifed
34	Fire
35	Explosion
77	All Other

Figure 1.2 shows the concept of variation of crash frequencies [Miaou et al. (1985)]. The total variation of crash frequencies can be decomposed into two components: systematic variation and random variation. To better understand this, it is assumed that the crash process could be repeated over and over again while keeping the five major factors constant for each site and time interval. The crash frequency for each site and time interval could be observed over and over again. The replication would allow the computation of the long-term mean value of crash frequency for each site and time interval. The variation of these mean values among those sites (between-site variation) and time intervals (between-time variation) is the systematic variation. The variation of crash frequencies observed from various replications about the long-term mean at each site and time interval is the random variation (within-site-within-time variation).

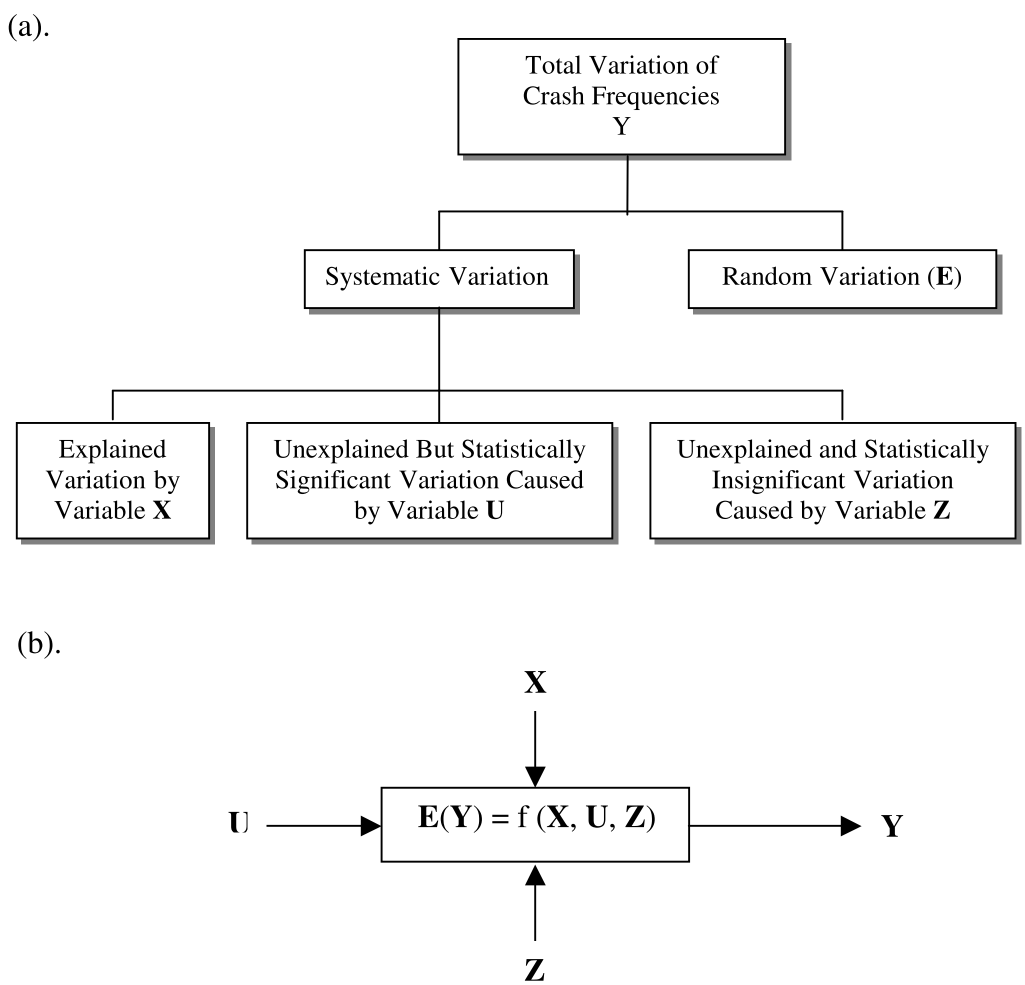

Figure 1.2. Variation Structure of Crash Frequencies [Miaou et al. (1985)]

The random variation can be thought of as the variation beyond explanation. Statistically, it is assumed that the random variation follows certain probability laws and can be characterized by the corresponding probability function, such as Poisson probability function and negative binomial probability function. The systematic variation is further decomposed into three types: (1) explained variation by variable X; (2) unexplained but statistically significant variation caused by omitted variable U; and (3) unexplained and statistically insignificant variation caused by variable Z. In reality, not all the information needed on the major crash related factors to explain the variation of crash frequencies is available. For example, in developing intersection crash prediction model, X may comprise traffic volume and some geometric parameters, which are used in the model, showing statistical significance in explaining the variation of crash. U may include vehicular parameters, driver information and on-site weather condition, etc. All of these factors are definitely crucial to the occurrence of crashes, but unfortunately are difficult to be included into the prediction model. Z consists of two types of factors: one, available but dropped due to statistical insignificance in explaining the variation, and another, unavailable and statistically insignificant.

Developing crash prediction models is a means of summarizing the complicated interactive effects of these crash related factors on the basis of information contained in the data, as well as engineering judgment (e.g. the selection of independent variables), and analytical assumptions about the crash process (e.g. which probability law will be relatively appropriate to apply to the crash study). A crash prediction model with good quality should estimate the occurrence of crash accurately at a specific statistical confidence level; meanwhile, the model shall make good engineering sense.

1.3.2. Poisson Regression Model

Many types of regression models have been used to develop crash prediction models in the past 30 years. However, conventional regression models are proved to be inappropriate by many studies [Jovanis et al. (1985), Hauer et al. (1988), Saccomanno et al. (1988), Miauo et al. (1993)]. Meanwhile, recent researches show that the Poisson regression model possesses the most desirable statistical properties in describing vehicle crash events that are random, discrete, nonnegative and typically sporadic.

Consider a set of n intersections of a given class (e.g. intersections with medium traffic volume, 50 mph posted speed on the major road, in the vicinity of business area, located in the urban area). Associated with each intersection i^{th}, is a set of parameters, X_{i1}, X_{i2}, …, X_{iq}, which describe the safety-related characteristics of this intersection, such as traffic volume, number of lanes on major road, posted speed on major road and/or minor road, etc.. Let the average number of crashes occurring at the i^{th} intersection during a specific time interval (e.g. crashes/per year or crashes/three-year) be denoted by Y_i, where, i= 1, 2, …, n. then denote the actual observation of Y_i during the same time interval by y_i, where y_i = 0, 1, 2, …. and i = 1, 2, …, n. The objective of a statistical model is to provide a relationship between a function of the expected number of crashes, $E(Y_i) = \mu_i$, at the i^{th} intersection, and the q parameters of this intersection, X_{i1}, X_{i2}, …, X_{iq}. This relationship can be formulated through a general linear form:

$$g(\mu_i) = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_q X_{iq} \tag{1-1}$$

where, the regression coefficients, β_0, β_1, β_2, … , β_q, are to be estimated from the data and the estimation procedure to be adopted is dependent on the assumption made about the distribution of Y_i. The assumption underlying Poisson regression is that the number of crashes, Y_i, follows a Poisson distribution with mean μ_i. The probability that an intersection defined by a set of explanatory variables, X_{i1}, X_{i2}, …, X_{iq}, experiences y_i crashes during a fixed time interval can be expressed as:

$$P(Y_i = y_i, \mu_i) = \frac{\mu_i^{y_i} \times e^{-\mu_i}}{y_i!} \tag{1-2}$$

where,

Y_i – discrete and random variable representing the number of crashes occurring at i^{th} intersection during a period of time;

y_i - actual or observed number of crash at i^{th} intersection during a period of time;

μ_i - expected number of crashes, the dependent variable corresponding to a set of predictor variables.

The natural logarithm link function is adopted in Poisson regression models.

$$\ln(\mu_i) = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_q X_{iq} \tag{1-3}$$

From mathematics perspective, it is not always clear in practice what link should be employed, and very often the data are analyzed by comparing several alternative choices. The reason to choose the natural logarithm link function here is taking into account the non-negative feature of crash count data. As stated by McCullagh and Nelder (1989), although canonical link may be found to be adequate over the range of the data, it is often dubious and logically unsatisfactory for extrapolation. By using natural logarithm link, $\ln(\mu_i)$ rather than μ_i obeys the linear model. This construction ensures that μ_i remains positive for all combination of independent variables and parameters. In addition, recent crash prediction studies also show that the natural logarithm link function is a reasonable choice. The Poisson probability function has only one parameter, mean, μ_i, and the variance, σ^2, equals the mean of the distribution. This inherent limitation of Poisson model is uncovered to be the major shortcoming of applying Poisson regression to crash prediction study.

Under the assumption of Poisson distribution, the regression coefficients, β_0, β_1, $\beta_2, \ldots$, β_q, are estimated by the maximum likelihood method. The likelihood function is the product of the individual probability density functions.

$$L(\mu(= \prod_{i=1}^{n} \frac{\mu_i^{y_i} \times e^{-\mu_i}}{y_i!} \tag{1-4}$$

This is a function of the parameter, μ_i, and through them, the parameters, β_0, β_1, $\beta_2, \ldots$, β_q, are estimated by maximizing the likelihood, or more usually, by maximizing the logarithm of the likelihood. Because the logarithm is a strictly monotone transformation, the values that maximize L will also maximize log-L, which can be written as,

$$LL(\mu) = \sum_{i=1}^{n} [y_i \ln(\mu_i) - \mu_i - \ln(y_i!)] \tag{1-5}$$

The actual maximization procedure always requires an iterative calculation.

1.3.3. Negative Binomial Model

Regarding the types of models used for crash frequency studies, Poisson regression models have been shown to be more appropriate than conventional linear regression models. However, the inability of the Poisson model to handle over-dispersed data is a major concern with regard to studying crash frequencies. This inability results from the major limitation of the Poisson regression model, which requires the variance of the dependent variable to be equal to its mean. Literature shows that most crash count data are likely to be over-dispersed, which means that the variance will likely be significantly greater than the mean [Shankar et al. (1995)]. When the mean and the variance of the data are not approximately equal, the variances of the estimated Poisson model coefficients tend to be underestimated and the coefficients themselves are biased.

This limitation can be readily overcome by using the negative binomial regression model, which assumes that the crash frequencies are distributed by negative binomial distribution. The negative binomial regression model is an extension of Poisson regression model and arises from Poisson regression model by adding an extra and independently distributed error term ε. For mathematics convenience the error term, exp(ε), is usually assumed to follow a gamma distribution with mean 1 and variance α. The resulting joint probability function, which is called negative binomial probability function, can be expressed as:

$$P\left(Y_i = y_i; \mu_i, \alpha\right) = \frac{\Gamma\left(y_i + \alpha^{-1}\right)}{\Gamma\left(\alpha^{-1}\right) y_i} \left(\frac{\alpha\mu_i}{1+\alpha\mu_i}\right)^{y_i} \left(\frac{1}{1+\alpha\mu_i}\right)^{\alpha^{-1}} \tag{1-6}$$

where,

Y_i - discrete, random variable representing the number of crashes occurring at i^{th} intersection during a period of time;

y_i - actual or observed number of crash at i^{th} intersection during a period of time;

μ_i - expected number of crashes, the dependent variable corresponding to a set of predictor variables;

α - dispersion parameter.

Note that the mean and variance of the negative binomial distribution of crash data can be expressed as:

$$E(Y_i) = \mu_i \tag{1-7}$$

$$Var(Y_i) = E(Y_i) * [1 + \alpha E(Y_i)] = \mu_i + \alpha \mu_i^2 \tag{1-8}$$

The second term on the right hand of equation (1-8), $\alpha\mu^2$, arises from the combination of the Poisson distribution with the gamma distribution assumption, and relaxes the constraints of Poisson distribution. Actually, as α goes to zero, the negative binomial regression yields the Poisson regression.

Like Poisson regression model, the relationship between the expected value of dependent variable and the corresponding q parameters, $X_{i1}, X_{i2}, \ldots, X_{iq}$, is still taken to be:

$$\ln(\mu_i) = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_q X_{iq} \tag{1-9}$$

The model coefficients, β_0, β_1, β_2, … , β_q, and the extra parameter, dispersion parameter α, are estimated by maximum likelihood method [Lawless (1987)]. The likelihood function is:

$$L(\mu_i, \alpha) = \prod_{i=1}^{n} \frac{\Gamma\left(y_i + \alpha^{-1}\right)}{\Gamma\left(\alpha^{-1}\right) y_i!} \left(\frac{\alpha\mu_i}{1+\alpha\mu_i}\right)^{y_i} \left(\frac{1}{1+\alpha\mu_i}\right)^{\alpha^{-1}} \tag{1-10}$$

The log-likelihood function is:

$$LL(\mu_i; \alpha = \sum_{i=1}^{n} \left[\ln\left[\frac{\Gamma(y_i + \alpha^{-1})}{\Gamma(\alpha^{-1}) y_i!}\right] + y_i \ln(\frac{\alpha\mu_i}{1+\alpha\mu_i}) + \alpha^{-1} \ln(\frac{1}{1+\alpha\mu_i}) \right] \tag{1-11}$$

All the estimation procedure of coefficients was done using the Statistical Analysis System (SAS).

1.3.4. Case-Based Crash Prediction

As mentioned previously, statistical approaches such as linear regression, Poisson regression and Negative Binomial regression analysis have been the main tool used to uncover the relationship between roadway parameters and crash frequencies. This research used these statistical approaches and also incorporated a new approach. Unlike the regression methods, this new approach is model free. It attempts to predict the crash frequency at a new intersection based on the (past and known) crash frequencies at similar intersections.

Since the new approach attempts to predict crash frequency at a new intersection based on some known cases, it is called case-based crash prediction (CBCP). Its foundations or origins rely on the early work by Schank and Abelson (1977) and Schank (1982). Its basic idea is to remember old solutions (crash frequencies) to similar problems (intersections) and to adapt them to fit a new problem (intersection) rather than having to solve it from scratch. In other words, CBCP requires access to past experience to improve system performance.

A CBCP system involves the following basic steps.

(1) Retrieve the most similar known cases (intersections) for application to the new case (intersection), with respect to roadway environment.

(2) Adapt or reuse the information and knowledge from the previous cases to solve the new case. The selected most similar cases have to be adapted when they do not match the new case perfectly.

(3) Evaluate the proposed solution (crash frequency) to the new case. A case-based reasoned requires some feedback to know what is going right and what is going wrong. Usually, this is done by performing some sort of search. In this study, the genetic algorithms (GA) approach is used.

Case-based reasoning approach has been successfully used in practice [Krovvidy et. al. (1993), Sanchez et al. (1997)]. Zhang and Yang's work (1997) on the main highways of Utah is the only literature available for traffic safety application. Experimental results by Zhang and Yang show that this approach is applicable to highway crash prediction and compared favorably with traditional methods in terms of prediction errors. Zhang and Yang did not, however, discuss the impact of traffic signal installation on crashes at intersections.

1.4. Research Objective

The main objective of this research was to develop statistical crash prediction models that can estimate the expected number of crashes at an intersection before and after the installation of traffic signals, in terms of total number of crashes and number of crashes for different crash types. The research mainly focused on the following objectives: (1) to estimate the change or impact of signalization on the expected number of crashes based on the expected total number of crashes and number of crashes for different crash types for the before and after period, (2) to use the Poisson regression and Negative binomial regression models in the determination of the prediction models, (3) to incorporate a new approach denominated case-based crash prediction, which is model free, in the evaluation of safety at intersections, (4) to explore the safety impacts of traffic signalization on intersection crashes on a statewide sampling of intersections in Florida through a before-and-after comparison analysis of yearly average number of crashes, crash rates, and crash severity, and (5) to evaluate the differences in the distribution of crashes by crash type, severity and surrounding land use type.

CHAPTER 2. LITERATURE REVIEW

2.1. Change of Crash Patterns at Intersections with Control Change

Crash patterns will change at intersections after the installation of traffic signals; however, the patterns changes are not clear. As stated by Box et al. (1970), the effect of installing traffic signals cannot be described specifically because the signal may reduce crashes under certain circumstances but widespread examples of higher rates after signal installation indicate the possibility of worse crash experiences under other circumstances. Regarding the effect of signalization on crash types and severities, it is indicated that right angle crashes tend to decrease, turning crashes and rear-end crashes tend to increase, and. the percent of injury crashes does not appear to increase.

King and Goldbatt (1975) carried out a comprehensive study to investigate the relationship of crash patterns to type of intersection control. They investigated 250 intersections located in nine different states (Colorado, Illinois, Maryland, Massachusetts, New York, Oklahoma, Pennsylvania, Washington and West Virginia). An analysis of crash data for the cases of before and after signalization was performed by using analysis of variance and regression techniques to show the relationship between crash patterns and type of control. The type of controls included signalization, four-way stop sign control, and two-way stop sign control. The study found that safety at an intersection was improved by the installation of a traffic signal, and that signalization leads to a significant reduction in right-angle crashes and a significant increase in rear-end crashes. The authors indicated that intersections after signalization may have higher crash rates that are usually offset by less disutility per crash, which leads to a no significant change in total crash-related disutility.

Short et al (1982) performed a before-and-after signalization crash study using 31 recently signalized intersections within the City of Milwaukee. Upon signalization, little or no change was noted overall either in the number of crashes, or in crash severities as measured by property-damage-only-equivalent (PDOE) and severity index (SI). A significant decrease of 34% in the number of right angle crashes and a significant increase of 37% in rear-end crashes after signalization were reported in the study. The authors also indicated a significance shift in crashes occurred between flash and normal operation time periods at thirty intersections that had flash operation at night. At these intersections, the PDOE crashes increased by 42% during flash operation, and decreased by 8% during normal operation. This shift was caused by a change in crash severity.

Shen (1984) performed an analysis to determine the safety aspects of newly signalized intersections in the District of Columbia. Twelve newly signalized intersections installed between January 1978 and June 1980 were randomly selected for the study. The author used the Poison distribution test to determine whether the changes in the number of crashes as a result of signalization were statistically significant. The conclusion of the research was that the decrease of approximately 35% of right angle crashes, 40% of sideswipe crashes, and 67% of pedestrian crashes could all be attributed to the installation of traffic signals. It was also found that the increase of more than 40% of rear-end crashes was statistically significant. Based on four of the twelve intersections, it was also indicated that when a signal was installed in conjunction with the opening of a new METRO station, it always resulted in higher crash frequencies even after signalization due to the substantial increase in traffic volume.

Radwan and Wing (1987) presented a report that contained a comprehensive review of signal installation and their impacts on crash patterns, crash frequency and crash severity. The report contained information on crash statistics by type and severity; crash rates for stop controlled and signalized intersections, crash patterns on arterials, and crash statistics for different signal types. Pedestrian safety due to signal installation was also addressed. This report indicated that due to signal installation, right-angle crashes decreased, rear-end and miscellaneous crashes increased, and overall crash rates did not change significantly. In regard to total number of crashes, no consensus in the results was presented. It was also concluded that in order to improve crash rates, intersections must have high traffic volumes, high existing crash rates, and complex geometric configurations before signalization becomes effective. The report also indicated that signalization is not a reliable measure for the reduction of crashes but it does not produce a significant increase in crashes either. Traffic signal removal was also addressed in terms of traffic safety. The analysis concluded that after signal removal the total number of crashes and injury crashes decreased; right angle crashes increased and rear-end crashes decreased. Finally, recommendations such as "Crash Reduction Using Signal Coordination on Arterial Streets" for future research were given.

Wattleworth et. al. (1988) developed several tables with estimated Florida crash reductions factors for different situations such as signalization, and channelization. For the signalization part, a total of ten improvements including new signal at channelized intersection, and new signal at non-channelized intersections were considered. For each improvement, the percentage of crash reduction for different crash types, crash severities, and weather conditions were listed. Their conclusions indicated that angle, left turn, and right turn crashes were reduced after adding traffic signals at channelized intersections. In reference to non-channelized intersections, only left-turn crashes decreased after a signal was installed. The specific changes of rear-end and other crash types were not analyzed in the study.

2.2. Traffic Crash Prediction Modeling

The most commonly used approach in the study of safety at intersections with new signals are the before and after signal installation studies. These studies can be applied to individual intersections where a new traffic signal has been installed. However, the conclusions for the particular intersection under study may not apply to other intersections of similar configuration and traffic demand. To overcome this problem, before and after studies may be performed collecting crash data at a large number of intersections in order to cover a wide range of traffic, roadway, and environmental characteristics. By using appropriate statistical techniques, such as analysis of variance and regression modeling, the safety impacts of traffic signal installation at intersections could be assessed.

Although before and after studies could provide some insights regarding the safety consequences of the installation of signals at intersections, engineers may be more interested in what factors cause these change, and the relationship between intersection crashes and these factors. Furthermore, it will be very helpful to be able to estimate the safety at intersections based on available crash related information. An appropriate approach to address these concerns is traffic crash modeling.

Up to date, most of the traffic crash modeling studies had focused in developing the statistical relationship between crashes at intersections or roadway sections and geometric variables. Then, the safety of an intersection can be estimated using the developed crash models [Hauer (1988), Bonneson et. al. (1993), Bauer (1996)]. However, very few studies had combined the before-and-after comparison approach together with the prediction modeling approach. Given some safety related explanatory variables, the combination of these two approaches gives the possibility for engineers to quantitatively estimate the safety at intersections before and after the installation of signals, and to investigate the impacts of signalization. Therefore, the crash prediction results from the before-after models would be a very helpful reference for decision-making.

David and Norman (1975) evaluated the relationship between motor-vehicle crash rates and geometric and traffic of a group of intersections with common design features. This analysis was based on a relatively detailed on-scene inventory of the geometry, design features, and traffic counts of 558 intersections coupled with police reports of 4372 crashes that occurred in those areas during the three-year study period. Forty-one crashes were investigated in-depth by a multidisciplinary team to determine causal factors and to evaluate the effects of federal safety standards on intersection crashes and severities, Finally, six design features including left-turn storage lanes and multiphase signalization were found to be crash-related-sight-distance obstructions.

Datta and Dutta (1990) worked on a research project to determine the changes in crash characteristics and crash severities at 102 newly signalized intersections in Michigan. It was also evaluated the effect of traffic signals at intersections with no geometric changes. This study used two to three years of crash data for both the before and after periods to evaluate the impact of signal installation based on crash rates. The conclusions based on statistically significant values of crash rates for the before and after periods at 67 intersections with no geometric improvements of the 102 locations were summarized as follows: total crash rate decreased by 15.5 percent; right-angle crash rate decreased by

52.5 percent; rear-end crash rate increased by 64.5 percent; head-on and left-turn crash rates increased by 75 percent; and other type crash rate decreased by 31.8 percent. In the study, there is no indication of the left turn treatment of the intersections considered but it is mentioned that only 28 percent of the locations had left turn lanes.

Huang and May (1991) employed a three level modeling approach for crash prediction models of unsignalized and signalized intersections. The objective of the study was the application of prediction models to understand the cause-and-effect relationship of crashes in unsignalized and signalized intersections. Their results indicated that for signalized intersections, actuated controllers were better from a safety point of view. For very wide intersections, multiphase actuated controllers were necessary in order to accommodate turning movements required in wide intersections. The study also indicated that the most important factors affecting crash occurrences in both unsignalized and signalized intersections were level of conflicts, where higher level of conflicts related to higher risk; and severity of conflicts, where higher speed related to more direct conflicts.

Bauer and Harwood (1996) worked in a research to develop statistical models of the relationship between traffic crashes and highway geometric elements for at-grade intersections. These models also incorporated the effect of traffic control features and traffic volumes on intersection crashes. The database used to develop the models was obtained from the California Department of Transportation. Field data were also collected for a sample of Urban, four-leg, signalized intersections to provide information on additional geometric design variables and turning-movement counts that were not available from the database. The statistical modeling approaches used in the research included Poisson, Lognormal, Negative Binomial, and Logistic regression, as well as discriminate and cluster analysis. Regression models of the relationships between crashes and intersection geometric design, traffic control, and traffic volume variables were found to explain between 16 and 38 percent of the variability in the crash data. However, most of that variability was explained by traffic volume variables considered; geometric design variables accounted for only a very small additional portion of the variability.

Vogt and Bared (1998) presented a report for the implementation of the Crash Analysis Module in the Interactive Highway Safety Design Model (IHSDM). This report described the collection, analysis, and modeling of crashes on rural roads in Minnesota (1985-1989) and Washington State (1993-1995). Poisson, Negative Binomial, and extended Negative Binomial models were developed. The models indicated that exposure and traffic counts were the key highway variables contributing to crashes. Also, it was found that the other variables affecting crashes at intersections were: vertical and horizontal alignments, roadside conditions, number of driveways, posted speed, approach angles, and turning lanes. In this study, advanced statistical techniques were applied to assess the explanatory value of the models in the presence of Poisson randomness and over dispersion. A non-parametric statistical modeling technique known as the Classification Regression Tree (CART) was used to group intersections by significance of prediction. This method seems to be very accurate, but it is also very complicated to apply.

2.3. Simple Linear Regression vs. Generalized Linear Regression

Researchers have attempted several statistical approaches when relating traffic safety measures (e.g. crash frequencies, severity-weighted crash frequencies, crash rates) to traffic related explanatory variables. Among them, simple linear regression and generalized linear regression are the two most commonly used statistical techniques to develop crash prediction models. Simple linear regression is the traditional approach to develop crash prediction models. In the classical linear model, the dependent variable (e.g. crash frequency) is expressed as a linear combination of explanatory parameters with or without interactions, under the assumption that the dependent variable is normally distributed. Unlike conventional simple linear regression, generalized linear models, such as Poisson regression, negative binomial regression and lognormal regression, are based on alternative distributions. Poisson regression is appropriate for dependent variables that have a Poisson distribution, as crash counts often do. Negative binomial regression assumes the negative binomial distribution, and lognormal regression assumes the lognormal distribution. For each of these models, the dependent variable can be crash frequencies, or similar safety measures mentioned above.

Even though simple linear regression has generated many useful findings, studies show that this approach suffers some undesirable statistical properties, for example, the poor explanatory ability of the variation in crash data. In the study performed by King (1975), the authors indicated that a linear regression model, even one with many independent variables, would not furnish an adequate model of crash experience associated with a given type of intersection control, and suggested to explore some more complex, probably non-linear regression model.

Joshua and Garber (1990) studied the relationship between crash involvements of trucks and associated traffic and geometric variables using both linear and Poisson regression models. The authors concluded that the multiple linear regression models did not adequately describe that relationship, but that the Poisson models did.

Miaou and Lum (1993) completed a study to evaluate the statistical properties of two conventional linear regression models and two Poisson regression models. The four types of models considered were: (1) an additive linear regression model; (2) a multiplicative linear regression model; (3) a multiplicative Poisson regression with exponential function; and (4) a multiplicative Poisson regression with non-exponential rate function. The authors concluded that of the four models tested, Poisson regression models outperformed linear regression models. Furthermore, the Poisson regression model with exponential rate function was the favored model.

Bauer and Harwood (1996) summarized several reasons indicating why conventional linear regression models are inappropriate for modeling crash frequencies or crash rates. The first reason indicate that traffic crashes are random and discrete events that are sporadic in nature. Secondly, crash frequencies for particular intersections or relatively small roadway sections are typically very small integers even if several years of crash data are obtained for those intersections and roadway sections. In fact, it is not uncommon for a substantial proportion of the sites in a crash study to have experienced no crashes at all during the study period. Small integer counts, often zero or close to zero, do not typically follow a normal distribution. Finally, crash frequencies and crash rates are necessarily non-negative, and traditional linear regression models could predict negative values for them.

2.4. Poisson Regression vs. Negative Binomial Regression

According to previous research, generalized linear regression definitely is a more adequate crash prediction approach than simple linear regression. Poisson regression models and negative binomial regression models are the generalized linear regression models that are being used widely. For the Poisson regression model, one important basic assumption is that the mean and the variance of the error distribution are equal. This feature simplifies the probability function, which only has one parameter. On the other hand, this advantage turns out to be the major disadvantage of Poisson regression models when applied to modeling crash data, which exhibits extra variation. If the variance of the crash frequencies exceeds the mean, then the data are over dispersed. When over dispersion exists in the data and Poisson regression models are used, the variances of the estimated model coefficients tend to be underestimated, which means the significance of the models will be overstated.

In their study, Miaou and Lum (1993) suggested the use of a more general probability distribution such as the negative binomial distribution to overcome the over-dispersion problem. In the follow-up study, Miaou (1994) recommended that the Poisson regression model should be used as the initial step to establish the relationship between the dependent variable and independent variables. Then, if over dispersion exists and is found to be moderate or high, both the negative binomial regression models and zero-inflated Poisson regression models can be explored.

J. Nicholson (1985) analyzed the considerable variation in the variability of crash counts. His results revealed that the pattern of crash occurrence at many locations was either too regular or too irregular to be well described by the Poisson process. Thus, the procedure for analyzing temporal variations in crash occurrences at particular locations should take into account the variations in the variability of crash counts. Based on the variance/mean ratio, the Binomial (variance/mean < 1.0) and Negative Binomial distribution (variance/mean ratio > 1.0) were complements to the Poisson distribution.

Poch and Mannering (1996) used seven years of crash data from 63 intersections in Bellevue, Washington, to estimate negative binomial regressions of the frequency of total crashes, rear-end crashes, angle crashes, and approach turn crashes at intersection approaches. The estimation results uncover the interactions between geometric and traffic-related elements and crash frequencies. In the study, each intersection was divided into separate approaches, and crash data were taken for each approach in one-year intervals. For the models, each intersection approach was considered as an observation, and a total of 64 possible explanatory variables were collected. The developed regression models identified significant traffic and geometric elements that tend to increase or decrease crash frequencies. The understanding of these elements can be beneficial to crash reduction at intersections.

Bernardo and Ivan (1997) utilized the Poisson regression to study the number of crashes versus crash rates at an unspecified number of intersections in Connecticut. The authors believed that the Poisson regression analysis was a better estimator of crashes than the linear regression analysis. In the study, three years of existing crash data were utilized for modeling. The intersections in the data were not separated into signalized or non-signalized locations, which have different impact on traffic operations. Results indicated that modeling crashes appears to be more logical than crash rates for the Poisson distribution, in the sense that the relationship between exposure and crashes is more accurate.

CHAPTER 3. RESEARCH APPROACH

3.1. Phase One – Crash Data Analysis

3.1.1. Methods of Evaluation

In order to conduct an objective before-and-after crash study, it is necessary to select an adequate crash database. FDOT has a very large crash database that is updated yearly. The database includes crashes gathered from the Department of Highway Safety and Motor Vehicles (DHSMV). Crash data maintained in the database are kept for five years. The crashes included in the database are those with a high amount of property damage, an injury, or a fatality. Crashes with high property damage are those with an estimated property damage of $500.00 or more. Crashes with minor property damage are not included in this database. The exclusion of crashes reported in short forms in the database may affect the estimated impacts of signalization on crashes in the sense that not all crashes occurred at intersections are considered, and specific type of crashes, such as rear-end crashes, may be under reported because many of these crashes have low property damage. For each crash, there are more than 300 variables used to describe the site and time of the crash, the geometric conditions, the traffic control, and drivers and pedestrian's characteristics. Among these variables, crash type, crash severity and surrounding land use type were used for the before-and-after comparison analysis. Other variables were incorporated to the modeling part of the project. Details of the database handling process are presented later in the data collection chapter.

3.1.1.1. Number of Crashes and Crash Rates

A summary of number of crashes for all crashes and by crash type can be used to identify the change of pattern of crashes after a signal is installed. As mentioned before, previous intersection crash studies [King et. al (1975), Short et. al. (1982), Shen (1984), Radwin et. al. (1987)] strongly indicate that adding a new traffic signal results in a reduction of right-angle crashes but in an increase in rear-end crashes. For this study, the yearly average number of crashes was considered when performing the before and after analysis for each intersection. Similarly, the yearly average number of crashes by type at each intersection was also used in the analysis. Table 3.1 presents the years considered for the average crash data according to different years of signal installation.

Because traffic volume of many intersections most likely changes after a signal is installed, crash rates are considered to be more accurate than the number of crashes. Therefore, in this research, the crash rates for all crashes and different crash types and crash severities were also calculated. Crash rate is defined as the number of crashes per million entering vehicles. The following equation is used to calculate crash rates at intersections:

$$\text{CRS} = \frac{1{,}000{,}000 \times A}{365 \times T \times ADT} \qquad (3\text{-}1)$$

where:

CRS = crash rate for spot;

A = number of crashes in this time frame;

T = time frame of the analysis, years;

ADT = average daily traffic volume.

Table 3.1.Yearly Average Crash Data for the Before-and-After Period

Year	Before	After
90	Crash data in 89	3 year average, 91,92, and 93
91	2 year average, 89, and 90	3 year average, 92,93, and 94
92	3 year average, 89,90, and 91	3 year average, 93,94, and 95
93	3 year average, 90,91, and 92	3 year average, 94,95, and 96
94	3 year average, 91,92, and 93	3 year average, 95,96, and 97
95	3 year average, 92,93, and 94	3 year average, 96,97, and 98
96	3 year average, 93,94, and 95	2 year average, 97, and 98
97	3 year average, 94,95, and 96	Crash data in 98

The procedure to compute crash rates for different crash types and crash severities is: (1) for each intersection, calculate the sum of crashes for all crashes and each crash type and crash severity for each year considered; (2) select the yearly average ADT for each intersection directly from the database; (3) calculate the crash rate for each intersection for each year considered; (4) average crash rates for each intersection are calculated, for a two or three year period based on years of crash data available at the intersection.

3.1.1.2. Types of Crashes

The installation of traffic signals has been found to influence crash patterns at intersections. The types of crashes commonly considered are: rear-end, right angle, left-turn, right-turn, sideswipe, and pedestrian-related. Among them, right-angle crashes have been found to have a significant decrease due to signal installation. Rear-end crashes will increase due to signalization. Left turn, right turn, and sideswipe crashes have both increased and decreased as a result of adding traffic signals according to previous studies. Crashes related to pedestrians are found to decrease after signal installation [King et. al (1975), Short et. al. (1982), Shen (1984), Radwin et. al. (1987)].

Table 3.2 lists all the different crash types maintained in the FDOT mainframe database. The crash types selected for mean value comparison are rear-end, angle, left-turn, right turn, sideswipe, and crashes related with pedestrian. Table 3.3 shows the crash types selected. All other crash types are added together in a category called "all other" crash type. In regard to the statistical analysis in the distribution fitting part, rear-end, angle, and left turn are investigated separately while right turn, sideswipe and crashes related to pedestrians are placed into the "all other" crash category due to the fact that each one of these crash types did not have enough number of observations for the analysis in the distribution fitting part.

3.1.1.3. Crash Severity

Although signalized intersections have been found to have a higher number of crashes and no significant decrease in crash rates after signal installation, it is believed that there is a reduction in crash severity. The reason for this relies on the fact that right angle crashes are reduced. An angle collision will usually have at least one injury, and it is also more common to have fatalities. Left turn and right turn crashes are similar to angle crashes. Therefore, the severity of these crashes is also similar to angle crashes. The coding scheme for the extent of injuries in FDOT database includes the following categories:

Table 3.2. Crash Types in FDOT Database

Code Number	Crash Type
1	Rear End
2	Head On
3	Angle
4	Left Turn
5	Right Turn
6	Sideswipe
7	Backed Into
8	Parked Car
9	w/Other Motor Vehicle on Road
10	Pedestrian
11	Bike
12	Bike in Bike Lane
13	Moped
14	Train
15	Animal
16	Sign/Sign Post
17	Utility/Light Pole
18	Guardrail
19	Fence
20	Concrete Barrier Wall
21	Bridge Abutment/Pier
22	Tree/Shrub
23	Construction Barricade/Sign
24	Traffic Gate
25	Crash Attenuators
26	Fixed Object Above Road
27	Other Fixed Object
28	Moveable Object on Road
29	Ran Into Ditch/Culvert
30	Ran Off Road Into Water
31	Overturned
32	Occupant Fell From Vehicle
33	Tractor Trailer Jack-knifed
34	Fire
35	Explosion
77	All Other

1. No Injury
2. Possible Injury: The person complained of pain or momentary loss of consciousness due to an injury during the crash, but no visible sign of injury is evident to the investigators.
3. Non-Incapacitating Injury: The person experienced a visible but not serious or incapacitating injury during the crash.

4. Incapacitating Injury: The person experienced serious, incapacitating, nonfatal injuries during the crash. Broken bones, massive losses of blood, or more serious injuries are rated in this category.
5. Fatality: The person died within 90 days of the crash as a direct result of injuries received during the crash.
6. Non-Traffic Fatality.

Table 3.3. Crash Type Categories for Mean Value Before-and-After Analysis

Crash type	Crash Code Number
Rear end	01
Angle	03
Left turn	04
Right turn	05
Sideswipe	06
Pedestrian	10
All others	02,07~09,11~35 &77

In the study, crash severity is categorized into three severity classes for the before and after analysis of mean values: fatal (F), personal injury (PI), and property damage (PD). The type of possible injury, non-incapacitating injury, and incapacitating injury categories are combined into a unique injury category. For the statistical analysis in the distribution fitting part, fatal crashes are combined with injury data, and only the categories property damage and injury are taken into account. The reason for this was the insufficient data available to perform a separate accurate distribution analysis for the fatal crash data.

3.1.1.4. Surrounding Land Use

The last method for the before-and-after analysis and distribution fitting is based on surrounding land use type classification. This primary classification is essential since urban and rural areas have fundamentally different characteristics, which significantly influence travel patterns, particularly those related to land use and population density. In this study, the impacts of surrounding land use type on number of crashes and crash rates were investigated. Also, different crash types in rural and urban areas were explored. For the distribution fitting part, only crashes under urban condition were analyzed because the number of crashes occurred in rural area was not enough to be divided by different crash types. The classification for urban and rural was taken directly from FDOT database. The code scheme for state roads in this database contained five categories as shown in Table 3.4. From coding description of urban and rural, the code number 1 and 2 are combined as Rural, and 3 and 4 are combined as Urban for the before and after analysis of mean values and distribution fitting. Originally, Outside-City (1 and 3) and Inside-City (2 and 4) were also investigated, but the results were not statistically significant and they are not shown in this report.

Table 3.4. Rural and Urban Categories

Code number	Description
1	Outside City, Outside Urban
2	Inside City, Outside Urban
3	Outside City, Inside Urban
4	Inside City, Inside Urban
5	Unknown

3.1.2. Statistical Analysis

3.1.2.1. Paired t-Test

Before and after analysis of mean number of crashes and crash rates on a statewide sampling of cash records were conducted to get the reduction or increase of crashes at intersections where traffic signals were being investigated to replace STOP/YIELD signs. Paired t-tests were conducted to determine if the difference between the before and after period was statistically significant.

Paired t-test is a special case of the two-sample t-test. It occurs when the observation on the two populations of interest are collected in pairs. Each pair of observations is taken under homogeneous conditions, which is at the same intersection in this study. Number of crashes and crash rates, both before and after signal installation for all locations, were compared by paired t-test to determine if there was a statistically significant difference between the two periods. There are two types of paired t-test that can be conducted: one-tail or two-tail. The one-tail test is used to test whether one mean is significantly greater than another. The two-tail test is used to test whether the means are significantly different. In this study, only two-tail tests are conducted to determine the change of all crashes and different crash types, severities, and surrounding land use after signalization.

The formula used for calculating the t statistic is shown here:

$$t = \frac{\overline{X}_b - \overline{X}_a}{S_d / \sqrt{N}} \qquad (3\text{-}2)$$

where:

$\overline{X}_b$ = the sample mean for 'before' case,

$\overline{X}_a$ = the sample mean for 'after case,

$$S_d^2 = S_b^2 + S_a^2 - 2\left[\frac{1}{N-1}\sum_{i=1}^{N}(X_{bi} - \overline{X}_b)(X_{ai} - \overline{X}_a)\right] \qquad (3\text{-}3)$$

S_b = the sample standard deviation for 'before' case,

S_a = the sample standard deviation for 'after' case,

N = sample size

If $t > t_{critical}$ ($t_{critical}$ is obtained from standard statistical tables), the difference in mean number of crashes and crash rates is statistically significant for an assumed level of significance α, where the degree of freedom is equal to the number of locations minus one. Therefore, the null hypothesis is rejected. The null hypothesis is that there is no significant difference between the mean number of crashes and crash rates for the "before" and "after" cases. A significance level α of 0.05 was used in the analysis.

3.1.2.2. Distribution Fitting

The average number of crashes and crash rates per year were calculated for each intersection with the use of SAS. Details of this procedure to process data will be explained in the next chapter. The estimated values are then plotted into histograms, where the independent variable (x-axis) is the average number of crashes per intersection and the dependent variable (y-axis) is the number of intersections. Poisson and Negative Binomial distributions are used to fit the frequency of crash data for the before and after period using the observed mean and variance. Subsequently, the Chi-Square goodness-of-fit test was used to test the hypothesis whether the number of crashes (or crash rates) follows a particular probability distribution. The following paragraphs present a brief introduction to Poisson and Negative Binomial distribution.

The definition of Poisson distribution is: if the mean number of counts (λ) in the interval is greater than zero (λ>0), the random variable X that equals the number of counts in the interval has a Poisson distribution with parameter λ, and the probability mass function of X is

$$f(x)=\frac{e^{-\lambda}\lambda^{x}}{x!}, \qquad x=0,1,2,\ldots.. \tag{3-4}$$

where,

λ -- observed mean value of the crash frequency

In regard to the negative binomial distribution, the probability function of X is:

$$f(x)=\binom{x-1}{r-1}p^{r}(1-p)^{x-r}, \qquad x = r, r+1, \ldots\ldots \tag{3-5}$$

where

r, p – two parameters calculated from observed mean and variance.

The mean and variance of this distribution of crash counts can be expressed in terms of parameters *p* and r as follows:

Mean = E(Y) = r/p (3-6)

Variance = Var(Y) = r(1-p)/p^2 (3-7)

3.1.2.3. The Chi-Square Test

The Chi-Square goodness-of-fit test is used to test the hypothesis whether the number of crashes (or crash rates) follows a particular probability distribution. The test procedure requires a set of randomly chosen samples of size n from X, whose probability density function is unknown. These n observations are then plotted into a frequency histogram of k class interviews.

O_i represents the observed frequency in the i^{th} class interval. The expected frequency in the i^{th} class interval denoted E_i could be calculated from the hypothesized probability distribution. The test statistic is,

$$\chi_0^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i} \tag{3-8}$$

where

O – observed frequency in the class interval i,

E – expected frequency in the class interval i.

It can be shown that, if the population follows the hypothesized distribution, χ_0^2 has, approximately a Chi-square distribution with *k-p*-1 degrees of freedom, where *p* represents the number of parameters of the hypothesized distribution estimated by sample statistics. This approximation improves as n increases. If the calculated value of the test statistic $\chi_0^2 > \chi_{\alpha,k-p-1}^2$, the hypothesis that the distribution of the population is the hypothesized distribution would be rejected. $\alpha = 0.05$.

3.2. Phase Two – Statistical Modeling

The second phase concentrated on developing statistical models that can estimate the average number of intersection-related crashes as well as the corresponding variances at an intersection, in terms of all and different crash types, before and after the installation of traffic signals. There are five cases considered: (1) total crash frequency (all crash types); (2) rear-end crash frequency; (3) angle crash frequency; (4) left-turn crash frequency; and (5) other crash frequency (including all of other crash types). For each case, two models were developed, one based on the data before signalization, the other based on the data after signalization. The reason to use two models for before and after situations respectively rather than one model (using a dummy variable to describe the before-after situations) is to reduce the time series effect of crash data.

The regression models adopted in this study are based on observed crash frequency distributions and previous researches. Two general types of statistical regression models have been considered to apply to the crash data: (1) conventional linear regression models; and (2) generalized linear models, including log-normal regression models, Poisson regression models and negative binomial regression models.

As mentioned before, many previous researches in this field show that conventional linear regression models are incapable to model the traffic crash data, which are non-negative, random, discrete and sporadic in nature. As alternatives, generalized linear models were explored and adopted in recent crash studies due to their advantages over conventional linear regression models.

3.2.1. Statistical Prediction Modeling Procedure

The crash modeling consists of seven major tasks: (1) to collect and reduce the crash data; (2) to analyze the crash data to determine the safety measures that were adopted as dependent variables in the modeling, and find appropriate probability functions to describe the random variation of crash frequencies; (3) to select and analyze the predictor variables; (4) to determine an appropriate functional form and parameterization, $f(.;\beta)$, to describe the effects of predictor variables on expected crash frequencies; (5) to estimate the regression parameters β in $f(.;\beta)$ using appropriate statistical algorithm based on crash data and probability assumptions; (6) to assess the quality of developed models, and make sure that the models make good engineering sense in addition to fulfilling statistical goodness-of-fit criteria; and (7) to apply the developed models, and convert the modeling results to tables for use. These tasks are briefly presented in the following paragraphs.

The modeling database was built by selecting the 518 newly signalized intersections collected from across the state from the crash database generated in phase one. This crash database generated in phase one was created from the Florida crash database maintained by FDOT, which consists of all crashes occurred on state roadways from 1989 to 1998. The 518 intersections included in the modeling database contained safety related characteristics and crash counts occurred within the influence area of those intersections. The process of generating the modeling database will be presented in detail later.

Based on data analyses, five types of intersection safety measures were adopted: (1) average number of all crashes per year; (2) average number of angle crashes per year; (3) average number of left-turn crashes per year; (4) average number of rear-end crashes per year; and (5) average number of all other crashes per year (including all the other crash types). For each safety measure, two models were developed as mention previously, the “before” and “after” models.

Another important issue was to determine which intersection characteristics should be used as predictor variables in the model. The principle to select the predictor variables was to include as many useful variables available in the FDOT database as possible. At the same time, the variables should be easy to obtain by FDOT traffic engineers when applying the models. According to these criteria, totally seven intersection characteristics including ADT of major road, urban/rural, land use of surrounding area, number of lanes on the major road, posted speed on the major road, type of median, and shoulder treatment were included in the model as predictor variables.

Based on crash frequency distributions and previous studies Poisson regression and negative binomial regression were chosen to estimate the model parameters. Generally, Poisson regressions can be used to build the relationships between crash frequencies and a set of predictor variables under assumptions that crash frequencies are Poisson distributed. However, Poisson regression has a limitation requiring the variance of the data to be equal to the mean. This restraint can be overcome by negative binomial regressions assuming crash frequencies are negative binomial distributed. Thus, for each model, Poisson regression was used as an initial step in the modeling process, with a negative binomial regression being applied where over-dispersion was founded existed in the crash data. Both in Poisson and negative binomial regressions, the regression parameters were estimated by maximum likelihood method with GENMOD procedure in SAS. Once the models were developed, two methods were applied to test the goodness-of-fit of the models: Pearson's R-square, and likelihood ratio index.

With the developed models, the expected number of all crashes and crashes by type at an intersection before and after signalization were estimated. Then, the changes of the estimated crash counts were calculated as the impacts of signalization. The calculated results were tabulated in order to furnish a simple and clear overview of the impacts of signalization on intersections with different characteristics.

3.2.2. Critical Issues with Crash Prediction Modeling

3.2.2.1. Dependent Variables

The dependent variables adopted in the crash modeling process include: (1) average number of all crashes per year before and after signalization, (2) average number of angle crashes per year before and after signalization, (3) average number of left-turn crashes per year before and after signalization, (4) average number of rear-end crashes per year before and after signalization, and (5) average number of all other crashes per year before and after signalization. The "all other" crashes includes all of the crashes except angle, left-turn, and rear-end crashes. The reason to use the average number of crashes per year is that the time frame used in this study is not uniform because of the limitation of the database. For example, for some intersections, three-year "before" or "after" crash data were available, for some intersections only one-year or two-year "before" or "after" crash data were available. Considering this fact, using the average number of crashes per year was the best choice.

Regarding the crash types, the selection was based on the results of data analyses. In addition to the average number of all crashes per year, the average number of angle, left-turn, and rear-end crashes per year were chosen as safety measures. All of the other types of crashes were aggregated in one category called the "all other" crash type and were also used as one of the safety measures due to insufficient crash counts for each one of these crash types at intersections.

After the dependent variables were determined, statistical distributions of the dependent variables were analyzed. It was found that the shapes of crash frequency distributions follow the Poisson distribution, which means that Poisson regression might be an appropriate choice in crash modeling. This confirms the results for distribution fitting from phase one.

3.2.2.2. Predictor Variables

Statistically, the more predictor variables in the model, the more predictive ability the model will have. Thus, the principle for selecting the predictor variables is to try to include as many predictor variables as possible, based on the data available and engineering judgment. The selection of predictor variables was incorporated in the database building process in the modeling process. For this selection of predictor variables, all possible factors that may affect the occurrence of crashes at intersections should be considered. These factors can be grouped into five categories: drivers, traffic, intersection or roadway segment, vehicles, and environment (e.g. weather condition) but, even though, four of the five factors play an important role in traffic safety, traffic engineers can only directly manage factors related to roadway through intersection design or improvement phases. Moreover, many of those factors cannot be adequately measure or control, such as driver's characteristics and reactions. Therefore, the variables considered for the models were basically in the intersection or roadway segment group. Furthermore, within this group only variables that were available in the database were considered due to the fact that the models should be easy to apply when evaluating an intersection. Finally eight predictor variables, such as ADT on the major road, number of lanes on the major road, posted speed, land use of surrounding area, and so on, were included in the final modeling database and were available for the crash predictive modeling. Among the predictor variables, ADT was transformed from continuous variable to categorical variable because the results of crash modeling were going to be tabulated. Other predictor variables were also categorized into different levels during the modeling process to generate the best modeling results.

The selected predictor variables were initially considered for all the models, including the models for all crashes and for each type of crash studied, even though, all crashes and each one of the different types of crashes have specific and particular factors affecting them. One of the reasons that supported this approach is the limited number of variables available in the database. It also has to be mention that the main purpose of the models is to evaluate the change of the number of crashes due to signalization and not to another specific factor.

3.2.3. Test of Over-Dispersion

Firstly, Poisson regression was performed during the modeling process for each case. After that, the crash data were tested for over-dispersion related to Poisson regression. If extra-Poisson variation is proved to be significant, the Poisson distribution assumption is violated; then the negative binomial regression model would be a more appropriate choice.

To test the over-dispersion of data, the mean deviance and Pearson's χ^2 ratio were used due to two reasons. First, these methods are widely used [Bauer et al. (1996), McCullagh and Nelder (1983)]; secondly, these methods are adopted by SAS software. Let L_s denote the maximum likelihood estimated from the saturated model that has as many parameters as observations, making each fitted value equal to the observed value, and let L_β denote the likelihood estimated by the current model. For Poisson regression model, log-likelihood can be expressed as,

$$\log(L_\beta) = \sum_{i=1}^{n} [-\mu_i + y_i \ln \mu_i - \ln(y_i!)] \tag{3-9}$$

$$\log(L_s) = \sum_{i=1}^{n} [-y_i + y_i \ln y_i - \ln(y_i!)] \tag{3-10}$$

where,

y_i - actual or observed number of crash at i^{th} intersection during a period of time;

μ_i - expected number of crashes, the dependent variable corresponding to a set of predictor variables.

The deviance, or G^2, is defined as minus twice the logarithm of the ratio of likelihood of the current model to the saturated model [Nelder et. al. (1972), Agresti (1990), Greene (1997)], and for Poisson regression, can be expressed as,

$$G^2 = 2\sum_{i=1}^{n}(y_i \ln \frac{y_i}{\mu_i}) \tag{3-11}$$

The deviance has an asymptotic distribution that is Chi-squared with degree of freedom equal to n-p, where n is the sample size and p is the number of parameters estimated. By forming the ratio of the deviance to its residual degree of freedom, n-p, an estimate of the scale constant $G^2/(n-p)$, called the mean deviance, can be found. For the Poisson regression, this scale constant should theoretically be equal to one. Values substantially in excess of one reflect over-dispersion of the data. The acceptable range for the mean deviance, $G^2/(n-p)$, is from 0.8 to 1.2.

Similar to the mean deviance statistic, the Pearson's χ^2 ratio statistic is also used to test the over-dispersion of crash data. The over-dispersion index can be calculated as,

$$\sigma_d = \frac{Pearson's\,\chi^2}{n-p} \tag{3-12}$$

where, n is the number of observations and p is the number of parameters used in the model. Pearson's χ^2 can be calculated by,

$$Pearson\chi^2 = \sum_{i=1}^{n} \frac{(y_i - \mu_i)^2}{Var(Y_i)} \tag{3-13}$$

where, for Poisson regression, $Var(Y_i)=\mu_i$. The value of σ_d tends to be one. If $\sigma_d>1.0$, then the data have greater dispersion than is explained by the Poisson distribution and a further analysis with a negative binomial error structure is required.

3.2.4. Evaluation of Goodness-of-fit of Models

So far there is no commonly acceptable measure that can give an absolute assessment of goodness-of-fit for generalized linear models. Therefore, several measures are selected and calculated, and jointly will give a relatively accurate evaluation of the models. First, deviance, as stated previously, is defined as minus twice the logarithm of the ratio of the maximum likelihood under current model and the maximum likelihood under saturated model. Thus, deviance describes lack of fit, greater deviance indicates poorer fit [Agresti (1990)]. Secondly, according to McCullagh and Nelder (1983), the Pearson's χ^2 is asymptotic to the χ^2 distribution with n-p-1 degrees of freedom for large sample sizes and exact for normally distributed error structures. Therefore, for a model, similar to deviance, the greater the Pearson's χ^2 , the poorer the fit. However, this statistic is not well defined in terms of minimum sample size when applied to non-normal distributions. Therefore, it should not be used as an absolute measure of model significance.

In traditional least square regression, the coefficient of determination, R^2, is frequently used to assess the goodness-of-fit of a model. It represents the proportion of variation in the data that is explained by the model. However, it was shown that R^2 is not an appropriate measure to assess the goodness-of-fit of crash prediction models due to their non-normal and nonlinear nature [Miaou et al. (1985)]. As a variation, a measure based on the standardized residuals, Pearson's R^2, can be calculated for each model to give some indication of the goodness-of-fit,

$$R_p^2 = 1 - \frac{\sum_{i=1}^{n} \frac{(y_i - \mu_i)^2}{\mu_i}}{\sum_{i=1}^{n} \frac{(y_i - \bar{y})^2}{\bar{y}}} \tag{3-14}$$

where,

R_p^2 -- Pearson's R-square statistic;

y_i -- observed number of crash at i^{th} intersection during a time period;

μ_i -- estimated number of crashes during a time period;

$\bar{y}$ -- average crash counts at all intersections of interest.

In addition, as the counterpart of R^2 in nonlinear regression, a measure of overall statistical fit, the likelihood ratio index can be computed as,

$$\rho^2 = 1 - \frac{L(\beta)}{L(0)} \tag{3-15}$$

where,

$L(\beta)$ --Log-likelihood at convergence;

$L(0)$ --restricted log-likelihood (all parameters are set to zero except for the intercept).

The value of 0.200 is quite satisfactory considering the variance in the data, and values tend to be generally lower than typical R^2 values [Ben-Akiva and Lerman (1985), Poch and Mannering (1996)].

3.2.5. Application of Crash Prediction Models

Once the parameters of crash predictive models were estimated, the average number of crashes in terms of all crashes and specific types of crashes before or after signalization can be estimated by replacing the regression parameters, β_0, β_1, β_2, … , β_q, with the estimated values, and the variables X_{i1}, X_{i2}, …, X_{iq}, with the corresponding values of the intersection characteristics. If a predictor variable is insignificant and was excluded from the final model, the variable would be omitted in the linear equation. However, the estimated average number of crashes will only provide a statistic of the safety measure either for an infinite number of intersections with the same characteristics or an intersection in an infinite time period with every characteristic unchanged.

3.3. Phase Three – Operational Research Modeling

3.3.1. Retrieve the Most Similar Intersections

In order to predict crashes and the impact of traffic signal installation at an intersection, it is necessary to see if there are similar intersections for which the impact due to traffic signal installation is known. Therefore, the term *similarity* should be defined. As mentioned earlier, each intersection will be characterized by seven variables. That is, each intersection can be represented by x = (x_1, x_2, x_3, x_4, x_5, x_6, x_7). Then, the similarity or closeness between two intersections x and y is defined as

$$distance(\mathrm{x}, \mathrm{y}) = \sqrt{\sum_{i=1}^{7} w_i \left(\frac{x_i - y_i}{\max_i - \min_i}\right)^2}\,, \qquad (3\text{-}16)$$

where w_i is the weight of the i-th variable to be determined, and $\max_i$ and $\min_i$ are the maximum and minimum values of the i-th variable. Function (3-16) is a normalized Euclidean measure between intersections x and y, and is called the distance function. Once intersection similarity is defined, the database is searched to identify a small number of intersections that is most similar to the new intersection (i.e., having the smallest distance function values).

In order to use the distance function (3-16), the weights w_i must be determined first. Many weight-learning methods are available [Wettschereck et. al. (1997)]. In this research, the genetic algorithms (GA) approach to learn the best weights is applied. Details are given later.

3.3.2. Adaptation or Reuse of Previous Known Cases

Let x be the new intersection whose crash frequency is to be determined by prediction. In addition, let y^1, y^2, …, y^k represent the k intersections identified from the database as described in last section. If f_j denotes the known actual crash frequency at the j-th identified intersection. Then, crash frequency $\hat{f}_x$ at the new intersection is predicted by

$$\hat{f}_x = \frac{\sum_{j=1}^{k} \rho_j f_j}{\sum_{j=1}^{k} \rho_j}, \tag{3-17}$$

where ρ_j is computed by $\sum_{i=1}^{7}(1 - \frac{\left| x_i - y_i^j \right|}{\max_i - \min_i})$. In this research, ρ_j is called the similarity function.

The number of intersections k selected from the database is yet to be determined. Several methods exist for finding the best k value [Duda et. al. (1973), Zhang et. al. (1997)]. In this study, k value would be determined by experiments, using a simple line search approach.

3.3.3. Evaluation of Crash Frequency Prediction

It is obvious that different intersections may be selected from the database if different weights w_i are used in the distance function (3-16). To evaluate what weights are the best, performance measures should be set. In this study, the squared error between actual crash frequency and predicted crash frequency was used as the measure of performance of CBCP. Particularly, the search of the best weights is conducted by GA.

The basic idea of GA can be simply described as: given a collection (population) of solutions (here, each solution represents a specific set of weights), GA seeks to "breed" good solutions by simulating the natural evolutionary process (survival of the fittest). To evaluate the goodness of a solution, a fitness function should be defined as follows.

$$F_y = (f_y - \hat{f}_y)^2, \tag{3-18}$$

where f_y and $\hat{f}_y$ are the actual and predicted crash frequencies at intersection y in testing database to be defined later. Let N be the total number of intersections in the testing database. Then, the average fitness over all the testing intersections is used as the final fitness of a solution (i.e., weights w_j). That is,

$$F(w) = \sum_{r=1}^{N} \frac{1}{N} F_r \, . \tag{3-19}$$

The best (fittest) solutions are allowed, by combining their best features, to breed new solutions in such a way that the population steadily improves (on the average) in terms of fitness function (3-19). In order to keep the population stable, the best solutions (including the newly produced solutions) survive into the nest generation while the worst solutions die off.

In this research, each solution is represented by a string of 28 binary variables with each weight defined by 4 binary variables. Therefore, each weight will have an integer value ranging from 0 to 15. Each binary variable is called a gene in GA terminology. A single point crossover is used to produce new solutions (weights). To prevent premature convergence at a local optimum, a mutation rate of 3% is applied to all new solutions. If a new solution is to be mutated, a gene is randomly selected to mutate.

The GA is executed for a fixed number of iterations or until the fitness function does not improve for a number of consecutive iterations, and the overall best solution (i.e., weights) is output as the final solution. The GA is summarized in the following paragraphs.

3.3.4. Genetic Algorithms

Step 1. Randomly generate the initial population of P solutions, and compute the fitness of each solution.

Step 2. Repeat for K iterations or until the best fitness of some weight meet the error tolerance (e.g., fitness < 0.05):

Select the best two unused solutions, and apply the crossover and mutation operators to breed P new solutions. Compute the fitness function value for each new solution. Let the P best of both the old and new solutions survive into the next generation.

3.3.5. Implementation

This section discusses the implementation of the CBCP described earlier. The collected intersections are partitioned into two sub-databases: training database and testing database. The training database consists of the majority intersections, and is used to train the system and predict intersection crash frequencies. The testing database consists of approximately 10% of the intersections and is used to evaluate the fitness of the weight solutions as described previously.

The computational experiments indicates that using a population size of P = 50 and K = 50 in the Genetic Algorithm is appropriate. A simple line search method is used to find the best k value (the number of intersections selected from the training database). In particular, the values 5, 8, 10, 12, 15, and 20 for k were tried, and found that $k = 10$ yields the best prediction. Therefore, k was set equal to 10 in the computational experiments.

CBCP can be used to predict total number of all crashes, angle crashes, rear-end crashes, left-turn crashes, and "all other" crashes, which includes the remainder of crashes, at intersections before and after traffic signal installation. CBCP can also be used to directly predict the impact of traffic signal installation on intersection crashes. But in this study, the impact is computed using the predicted crash frequencies before and after signalization.

Since CBCP is based on known data, its prediction accuracy is closely dependent on the availability of data. With all the seven characteristic variables that define an individual intersection, there will be a total of 192 different intersection types. This leads to, on the average, fewer than 3 data points for each unique intersection type. This can make the prediction results difficult to use. To overcome this, statistical regression analysis using the predicted crashes was applied

3.4 Summary

The research approach used in each one of the three phases of the project was presented in this chapter. For the crash analysis, the methodology used in identifying the critical group of intersections needed for the analysis and the evaluation method used for the analysis and distribution fitting were described. In the statistical analysis, a paired T-test, the Poisson and Negative Binomial distributions were implemented. In reference to the statistical crash predictive modeling, once the statistical properties of crash frequencies were explored to determine the best regression model to use, the modeling procedure was presented with a detail description of each one of its steps. The steps included: analysis of dependent variables, analysis of predictor variables, Poisson regression and negative binomial regression, test of over-dispersion, evaluation of the goodness-of-fit of developed models, and applications of modeling results. Finally, the methodology used for the crash-based prediction model was presented. The steps followed for this methodology included the retrieve of most similar intersections, reused of previous known cases, evaluation of crash frequency prediction and implementation of the procedure.

CHAPTER 4. DATA COLLECTION AND REDUCTION

4.1. FDOT Crash Database

FDOT maintains a very large crash database generated by merging crash data from the Department of Highway Safety and Motor Vehicles (DHSMV) with roadway information from FDOT. This database is updated yearly. All reported crashes with a fatality, an injury, and high property damage occurred on state roads are included in this database. Therefore, the FDOT database not only contains crashes occurred at intersections, but also crashes occurred on roadway sections.

4.1.1. Crash Data Format

The crash data used in this study cover a 10-year period from 1989 to 1998. Corresponding to each year, there is one data file consisting of all crashes occurred on state roads during that year. For each crash, several record types containing specific information related to the crash are included. Table 4.1 shows the different record types for each crash. All ten files, stored in ASCII format, have the same database structure. As an example in ASCII format, the first two numbers “00” indicate the record type, the next 8 numbers represents the crash number; the twelfth number is the district number, and so on. A SAS program was written and used in order to change the ASCII format to SAS format.

4.1.2. Format Change Using a SAS Program

Several factors were considered in order to read the original database and to create the variables in SAS. One of them refers to the number of lines for record types "02" and "04", which may be different, and the varied length of each record type from "00" to "11". Additionally, 168 variables were selected for the original database for the research. The variables were selected based on their possible contribution to crash occurrences. These variables were selected from five of the twelve record types. The variable selection process is explained in detail later in the chapter. The record types selected were record "00"(Time and Location), record "01"(Characteristics), record "09" (RCI-Features-I), record "10"(RCI-Features-II), and record "11"(RCI-Point). In order to put the 168 variables in one file, these files with record type "00", "01", "09", "10" and "11" were merged into one merged file for each year.

Table 4.1. Description of Record Type

Record Type	Description
00	Time and location
01	Characteristics
02	Vehicle
03	Towed
04	Driver
05	Passenger
06	Pedestrian
07	Property Damage Amount
08	Reserved for future use
09	RCI–Features-I
10	RCI-Features-II
11	RCI-Point
12	RCI-Total

4.1.3. Code Check

The code for different crash types was changed in the FDOT crash database in 1993. Therefore, a code check was required for crash data files from 1989 to 1992. FDOT personnel were certain that the code for crash types for 1991 and 1992 were updated when the code was changed in 1993, but they were not certain about 1989 and 1990. The code change means that the numbering scheme for the different types of crashes changed in 1993. In the scheme, the codes used to define a crash type are numeric. As an example for the code change, an "angle" crash is coded as "03" since 1993 but it was coded as "02" before 1993. Table 4.2 shows the difference between the new code and the old code for the first ten numbers in the scheme of crash type. In order to check if the data for 1989 and 1990 had the new or old code in the FDOT database, a random number of crashes were selected for each one of the years. For each one of the crashes, the information was checked and compared with the information for the same crashes pulled out from the Fatality Analysis Reporting Systems database. This database has the old code for the crash data up to 1992. The comparison was done or match base on the crash number, which is unique for any crash and it is listed in both databases. Once the code was checked based on the comparison, it was found that the code from 1989 to 1992 had been already updated to the new code in the FDOT crash database. Therefore, no code adjustment was necessary for the yearly database utilized in this project.

Table 4.2. Examples of Codes for Different Crash Types

1989-1992	1993	Code Explanation
00		Not Applicable
01	01	Collision with MV in Transport (Rear-end)
02	03	Collision with MV in Transport (Angle)
03	04	Collision with MV in Transport (Left-turn)
04	08	Collision with Parked Car
05	06	Collision with MV in Transport (Sideswipe)
06	07	Collision with MV in Transport (Backed Into)
07	05	Collision with MV in Transport (Right Turn)
08	27	MV Hit Other Fixed Object
09	17	MV Hit Utility Pole/Light Pole
10	02	Collision with MV in Transport (Head-on)

4.2. Intersection Sample

The principle for collecting the intersection sample was to obtain as many intersections as possible to better evaluate the impacts of signalization on intersection crash experience. For this reason, seven FDOT district offices were contacted and with their cooperation 518 intersections were identified. These intersections had traffic signals installed during the period from 1990 to 1997. Furthermore, almost all the intersections collected were considered in the research due to the limited number of intersections available, and only intersections with major improvements besides signalization were eliminated. No random selection of intersections was performed, which raises a concern in regard to the possibility of a bias on the intersection sample. But, since this research is evaluating the impacts of signalization on intersection crashes and the sample size was limited, the sample selected consisted of almost all the newly signalized intersections. The signalized intersections are located on state roads with two to seven lanes on the major road in rural and urban areas. Table 4.3 shows the intersection sample collected from FDOT district offices by district and activation year. The source of the intersection sample shows that this sample is a representative at state level. Table 4.4 shows the different districts of FDOT.

4.3. Signal Activation Date

There are two very important dates regarding the installation of a traffic signal. The first is the maintenance agreement date and the second is the actual date when the traffic signal is activated. In this study, the activation date was used. The data from the signal activation year was not considered for the analysis. Details are presented in the following subsection.

Table 4.3. Intersections by Location and Activation Year

District	Activation Year								Total
	1990	1991	1992	1993	1994	1995	1996	1997	
1	6	9	2	7	6	0	5	5	40
2	12	9	8	12	11	6	4	4	66
3	6	12	6	10	11	10	14	13	82
4	1	6	11	11	0	5	0	0	34
5	36	29	21	31	21	16	14	11	179
6	30	0	0	0	30	4	5	26	95
7	0	0	1	1	6	5	5	4	22
Total	91	65	49	72	85	46	47	63	518

Table 4.4. Districts in FDOT

District Number	Location of District Office	County of District Office
One	Bartow	Polk
Two	Lake City	Columbia
Three	Chipley	Washington
Four	Fort Lauderdale	Broward
Five	Deland	Volusia
Six	Miami	Dade
Seven	Tampa	Hillsborough
Eight	Turnpike District	Orange

4.4. Time Frames

The crash data utilized for this study from the original FDOT crash database must be isolated from the remainder of the data file. One of the first steps in this data reduction process was to choose analysis time frames and discard data from outside the time frames.

Regarding time frames for crash analysis, the *Manual of Transportation Engineering Studies* indicates that a three-year window is the most common choice [Hummer (1994)]. Choosing a three-year time frame has several advantages. A three-year time frame enables analysts to collect sufficient crash counts. Whether or not enough crash counts are available is always the major concern for the analysts when conducting traffic crash studies, because traffic crash events are sporadic in nature, a large proportion of intersections could experience no crash at all if the time window selected is too narrow. In addition to aiding to increase the crash sample size, to some extent a three-year time frame could reduce the regression-to-the-mean effect compared to a shorter time frame. Regression-to-the-mean is a common phenomenon encountered in crash analysis studies. Simply stated, it means that the occurrences of crashes at an intersection vary statistically from year to year even if the conditions of the intersection have not changed, and have the tendency to regress to the long-term mean value. For example, the average number of crashes at an intersection during a long-term period is assumed to be M. If the number of crashes in the first year is $M_{1,}$ which is larger than M, statistically the probability of having less crashes than M_1 in the second year ($M_2 < M_1$) will be higher than having more crashes than M_1 ($M_2 > M_1$) under the same conditions. So, statistically the average number of crashes during a three-year period will be closer to M than the average number of crashes during a one-year period. Considering this, a three-year time frame will enable the crash data to represent the real pattern of intersection safety more accurately than a shorter time frame. Also, compared to a four-year or five-year time frame, a three-year time window is not so wide that some changes in background conditions can be tolerated within the scope of the study. Thus, a three-year time frame represents a good compromise between the desire for larger crash sample size and the desire for time frames within which conditions were unlikely to have changed a great deal. Therefore, a three-year time frame was used in this study. For example, for an intersection where signals were installed in 1994, crash data for 1991, 1992, and 1993 were kept in the database for the "before" period, and crash data for 1995, 1996, and 1997 were kept in the database for the "after" period.

Due to the fact that the original database covers a ten-year period from 1989 to 1998, some intersections have only one or two-year crash data available for the "before" or "after" period. In this case, time frames had to change according to the data available. For example, for those intersections where signals were installed in 1990, the time frame for the "before" study is one-year because only the crash data for 1989 are available in the original database. But the time frame for the "after" study is still a three-year period because the crash data for 1991, 1992 and 1993 are available in the database. Similarly, for those intersections where signals were installed in 1996, the time frame for the "before" study is a three-year period, which means crash data for 1993, 1994 and 1995 were included in the database for the "before" study. But the time frame for the "after" study is two-years, because only crash data for 1997 and 1998 are available in the original database.

Table 4.5 shows the time frames for intersections where traffic signals were installed in different years. The white cells filled with "before" represent the years in which crash data were kept for the "before" study. The white cells filled with "after" represent the years in which crash data were kept for the "after" study. The blank shaded cells represent the years in which crash data will not be used in the study. Table 4.6 presents the time frame information in numbers. This table also includes the number of newly signalized intersections gather per year.

Table 4.5. Time Frame for Intersections with Signals Activated in Different Years

1989	1990	1991	1992	1993	1994	1995	1996	1997	1998
before	1990	after	after	after					
before	before	1991	after	after	After				
before	before	before	1992	after	After	after			
	before	before	before	1993	after	after	after		
		before	before	before	1994	after	after	after	
			before	before	before	1995	after	after	after
				before	before	before	1996	after	after
					before	before	before	1997	after

Table 4.6. Number of Newly Signalized Intersections per Year

Year	Before (years)	After (years)	Number of Intersections
90	1	3	91
91	2	3	65
92	3	3	50
93	3	3	72
94	3	3	85
95	3	3	46
96	3	2	47
97	3	1	46

In Table 4.5 the cells filled with the figures "1990" - "1997" are also shaded. That means that crash data within the year when the signals were installed were not used neither in the "before" study nor in the "after" study. One of the reasons for this decision considers that it takes time for drivers to get used to the new signals. In other words, during the time period right after signalization, the driver behavior may be affected to some extent and the occurrence of crashes may appear to be abnormal. In addition, in the modeling process the average number of crashes per year based on one-year, two-year or three-year periods was used as dependent variable. If the crash data for several-months before or after the installation of traffic signals were used, data would have to be converted into annual average number of crashes. This type of data conversion is not commonly used in statistical analysis. Therefore, discarding the crash data for the activation year makes the dependent variable more consistent and easier to handle.

4.5. Identification of Intersection-related Crashes

The roadway numbering and milepost systems were used to identify from the FDOT crash database a crash occurred within the influence area of an intersection. Within FDOT, every state road has been given an eight-digit code called "Section" number that uniquely defines that roadway. The first two digits are the county number, the next three digits are the actual section number of the roadway, and the last three digits are known as the subsection number. In addition to this numbering system, a milepost system is being used to label a point located on a roadway. Most state roads in the State of Florida are labeled either from south to north or west to east. Milepost zero begins at the southern or western most terminus of the road within that county. Mileposts are kept with three digits after the decimal point. Thus, any point within the right-of-way of a roadway will be accurately identified using the numbering and milepost systems. For example, an intersection will be uniquely pinpointed once its eight-digit "Section" number and the milepost of its center point are available.

The roadway numbering and milepost systems are also applied in the FDOT crash database to identify the exact site location of a crash. There are totally five variables used to convey the information in the database: DISTID, COUNTYID, SECID, SUBSECID, and MILEPOST. The first four variables are used to identify on which road a crash occurred. MILEPOST is used to locate the exact position where the crash occurred. If the crash vehicle ran off the road, MILEPOST records the milepost of the point on the roadway that is nearest to the crash site. The subsection number is only used when a roadway is reconfigured (a one way pair is constructed which used to be a four-lane roadway). Consequently, a subsection is usually "000", but occasionally it could be another number such as "001" if reconfigured. Table 4.7 shows the format for the section number of the state road system.

Table 4.7. Format of Intersections from FDOT Districts

Section No.	Mile Post	Signal Activation Date
16006000	2.484	03/22/91
16006000	3.239	03/22/91
16006000	3.74	03/22/91
16011000	2.762	03/12/96
16011000	5.379	02/12/90
16060000	11.871	08/02/94
16110000	5.836	02/23/93
16110000	21.354	02/17/93
16140000	0.373	02/12/90

Based on the numbering and milepost systems, all crashes occurred within the influence area of an intersection can be identified by searching the FDOT database by the crash ID number and milepost range according to the ID number and milepost of the intersection. The influence area of an intersection considered in the research to investigate the impacts of signalization on intersection crashes is defined as a distance (250 ft) from the center point of the intersection in either direction of travel along the major road. This distance could be shorter in some special cases such as intersections closely spaced. The geometric information of the minor road for each intersection is not available. Thus, all crashes occurred within this 500 ft on the major road would be identified as intersection-related crashes. In other words, the influence area covers a 500 ft long area within the right-of-way of the roadway, and the base point to identify crashes within the intersection is its center point. This method is also used by other intersection crash studies [Bauer, et. al. (1996), Ogden et. al. (1996), Sayed et. al. (1999)]. Figure 4.1 shows the concept to identify intersection-related crashes.

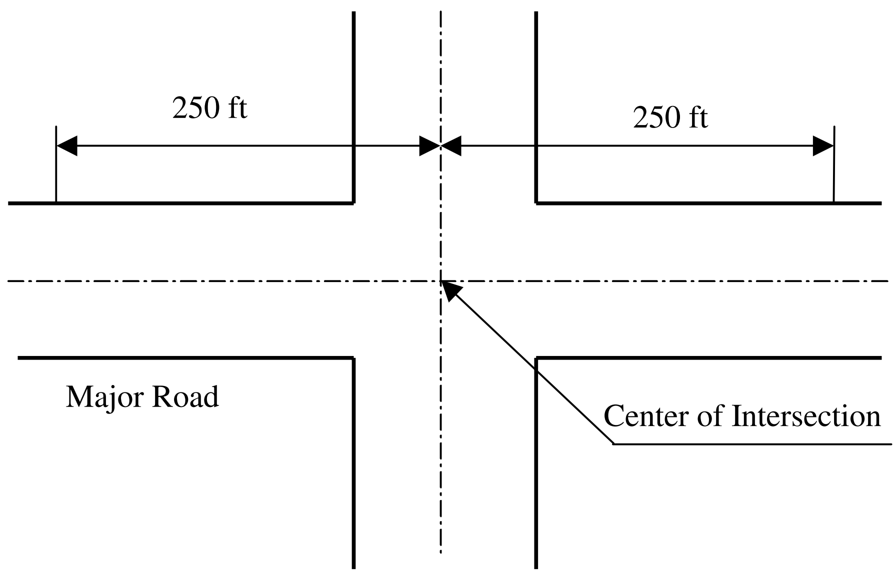

Figure 4.1. Concept to Identify Intersection-Related Crashes

4.6. Data Reduction Using SAS Programming

4.6.1. Software and Basic Concepts

For the project, the crash database containing only crashes for the newly signalized intersections was needed. For this purpose, a SAS program with Structured Query Language (SQL) was written to automatically gather all of crash data needed from the merged files. The SQL procedure implements Structured Query Language (SQL) for the SAS System. SQL is a standardized, widely used language that retrieves and updates data in tables and views based on those tables. The SQL procedure in SAS allows to:

- Retrieve and manipulate data stored in tables or views.
- Create tables, views, and indexes on columns in tables.
- Create SAS macro variables that contain values from rows in a query's result
- Add or modify the data values in a table's columns or insert and delete rows.

In this study, each intersection in the list needed to be matched with several different files. As mentioned before, if the signal was installed in 1991 at an intersection, crashes occurred in the before period need to be gathered from the files that contain crash data for 1989 and 1990, and in the after period from crash data files for 1992, 1993 and 1994. If the intersection had been signalized in 1996, crash data from 1993, 1994, and 1995 were needed for the "before" period, and crash data from 1997 and 1998 were needed for the "after" period. Figure 4.2 illustrates the relationship between the ten FDOT crash data files, the intersection list and the final "before" and "after" files. After the matching, the result file for “before" contains all the crashes occurred at the newly signalized intersections for the time period between 1 to 3 years before the signal was installed, and the "after" file contains the crash information for the period of 1 to 3 years after the signal was activated.

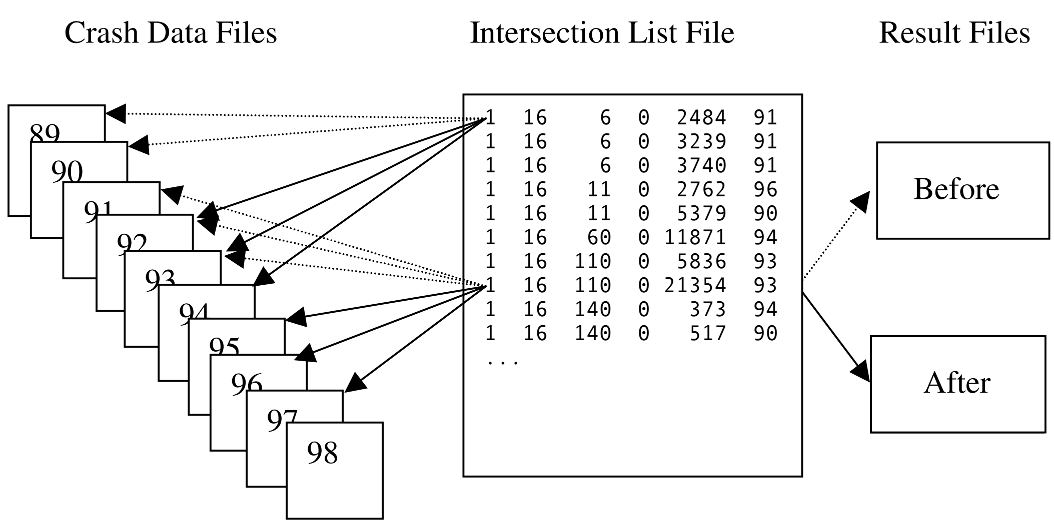

Figure 4.2. Relationship Among Files

4.6.2. SAS Programs and Flow Chart

In order to get the crashes from the yearly-merged-files, each intersection in the list of newly signalized intersections needed to be matched with each yearly-merged-file. For example, for intersections with signal installed in 1991, 1992 and 1993, the yearly merged crash data file "94" contains the crashes to be taken into the final "after" file. On the other hand, for those intersections where a signal was installed in 1995, 1996, and 1997, the crashes contain in the yearly merged crash data file “94” were taken into the final "before" file. This is the basic structure in the SAS program.

Figure 4.3 shows the procedure used to get the crashes for the "before" period. In the program, the ‘before’ file was established by matching the crash file in 1996 with the intersections with signal activated in 1997. Then, crash data from 1995 matching the intersections with signal activated in 1996 or 1997 were added to the “before” file. The “before” file was completed applying the loop method to the same procedure for the data from yearly-merged-files from 1989 to 1994. Each yearly-merged-file had three years of intersections with signal activation from (i+1) to (i+3), where "i" is the year of the crash data.

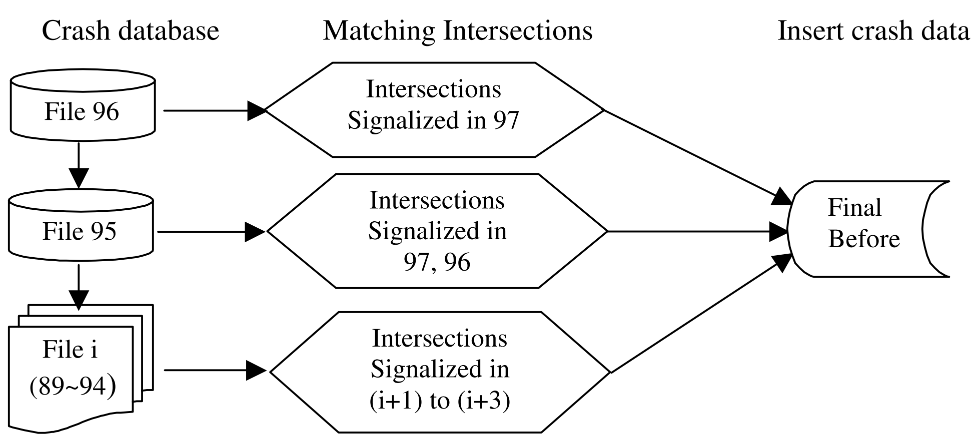

Figure 4.3. Data Matching Procedure for "Before"

Similarly, Figure 4.4 shows the procedure to get the "after" file. This file can be established by matching the year 1991 crash file with the intersections where the stop sign was changed to signal in 1990. Then, crashes from 1992 matching intersections where the signal was installed in 1990 and 1991 were added to the “after” file. Finally, the “after” file was completed by adding crashes from each crash data file from 1993 to 1998 matching three years of intersections from the year (i-3) to (i-1), where "i" is the year of the crash data. Two SAS data sets with 4565 and 6122 crashes occurred at the newly signalized intersections for the before and after period respectively were obtained

4.7. Preliminary Screening of Variables

As mentioned earlier, there are 12 record types for each crash to describe the crash-related information in the FDOT crash database. Each data record consists of a fixed number of variables. For example, there are totally 49 variables in record type 00, and only 9 variables in record type 03. At most, there will be up to 295 variables to describe a crash. Appendix A lists all the variables in the database.

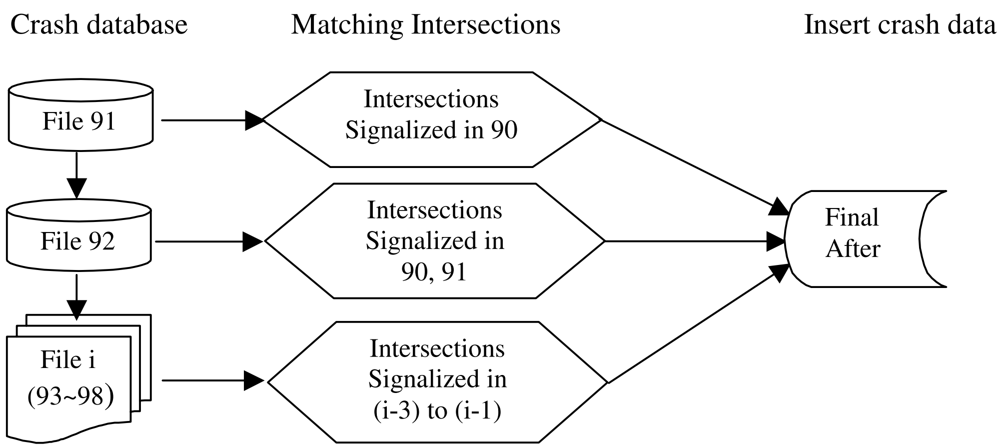

Figure 4.4. Data Matching Procedure for "After"

One step necessary in the database processing was to drop some variables that are not useful for this study in order to make the database smaller and easier to manipulate. First, several types of variables, including driver-specific variables, passenger-specific variables, pedestrian-specific variables, vehicle-specific variables, and time-specific variables were dropped. For an individual crash, these variables definitely are crucial to explain how the crash occurred. However, these variables are not useful in developing the aggregate relationship between the average number of crashes occurred at an intersection and the safety-related characteristics of this intersection. Based on this criterion, record types 03, 04, 05, 06, 07 and 12 were discarded. In record types 00, 01, 02, 09, 10, and 11, many variables that provide useful information for the study were kept. Appendix B lists the 62 variables that were included in the database after the preliminary screening.

4.8. Further Screening of Variables

During the data analyses, it was found that some variables were not necessary for the study due to different reasons. The variables eliminated from the database are classified into the following categories and are presented one by one in the following subsections.

4.8.1. Variables Describing Severity of Individual Crashes

Eight variables describing the severity of individual crashes were discarded because crash severity wee not modeled in this study. These variables are presented in Table 4.8.

Table 4.8. Crash Severity Variables

No.	Variable	Description
1	ACCSEVER	Crash Severity
2	INJURSEV	Injury Severity
3	DAMAGSEV	Damage Severity
4	DAMAGAMT	Total Damage Amount
5	VEHDAMAG	Total Vehicle Damage Amount
6	PROPDAM	Total Property Damage Amount
7	TOTFATAL	Total Fatalities
8	TOTINJUR	Total Injuries

4.8.2. Variables with High Proportion of Missing Values

Among 4565 crashes in the “before” data, there are 3411 crashes (74.72%) with missing values for the variables shown in Table 4.9. Among 6122 crashes in the “after” data file, there are 4640 crashes (75.79%) with missing values for those variables. The missing value is denoted by a symbol " . ", which means that the variables were not given any value while inputting data from hardcopy crash reports to compute-based database. These variables were eliminated from the database.

Table 4.9. Variables with High Percentage of Missing Values

No.	Variable	Description
1	ACCESCTR	Access Control Type
2	PREVLAND	Prevailing Land Use
3	HWYSHTP1	Highway Shoulder Type
4	HWYSHW1	Highway Shoulder Width
5	MEDIANW	Median Width
6	MEDIANTP	Median Type
7	HORPTINT	Horizontal PT of Intersection
8	SUPERELV	Super elevation
9	VERTPTIN	Vertical PT of Intersection

4.8.3. Variables with High Percentage of Unknown Values

In addition to variables with missing values, a group of variables were found to have very high percentage (even 100%) of unknown values according to each variable's coding system. The unknown values are denoted by "0", "9", "99", "unknown", or other symbols. Table 4.10 lists these variables and their percentage of unknown values.

Table 4.10. Variables with High Percentage of Unknown Values

No.	Variable	Description	Crashes with Unknown Values	
			"Before"	"After"
1	CROSSTRF	Cross Traffic	100.00%	100.00%
2	PASSDIST	Passing Distance	81.01%	100.00%
3	RDSONSIS	Roadway Consistency	81.01%	100.00%
4	RDALIGN	Roadway Alignment	81.01%	100.00%
5	STOPDIST	Stopping Distance	81.01%	100.00%
6	POINTADT	Point ADT	100.00%	100.00%
7	TPINTER	Type of Intersection	100.00%	100.00%

4.8.4. Variables Providing Similar Information

Some variables describe similar characteristics of the intersection. In most cases, the information the variables provide is consistent between them although the values of these variables are different due to different codes. In each case, one of the similar variables was kept in the database while the other(s) was dropped after careful examination and comparison. This method applied to the following variables:

(1) NUMBLANE (DOT number of lanes) and NUMLANES (DHSMV number of lanes). Both of them describe the number of lanes on the major road but from different sources. Either one of them can be used in the database. NUMLANES was chosen.

(2) RURURB (DOT rural/urban) and URBRUR (DHSMV rural/urban). Both variables describe the type of land surrounging the intersection. URBRUR was chosen.

(3) Totally 4 variables were used to describe the crash-related speed. ESTSPEED describes the estimated speed of the first vehicle involved in each individual crash; POSTSPED describes the posted speed on the major road; MAXPSTSP describes the maximum posted speed; MINPSTSP describes the minimum posted speed. The first speed variable was eliminated because it only describes individual crashes. MINPSTSP was dropped due to the fact that its vlaue was zero for more than 98% of total crashes. Either POSTSPED or MAXPSTSP could be used in the modeling. In the project, POSTSPED was chosen.

(4) ROADTYPE and DIVIDNOT are used to describe whether the major road is divided or not. DIVIDNOT was chosen.

(5) CLASSCAT determines the class/category of the roadway, providing the comprehensive information provided by DIVIDNOT, NUMLANES, and URBRUR. This variable was discarded for the convenience of data process.

4.8.5. Other Discarded Variables

Several other variables were eliminated from the database. The following variables were eliminated because their limited effects on intersection safety. ROADSYS describes the type of function of a roadway where a crash occurred, such as interstate highway, state road, local road, or turnpike. ROADSURF, PAVINDEX and PVSURFTP are used to describe the type of pavement. SITELOC and SITLOCAT are used to describe the site location of individual crash, for example, at intersection, on bridge, on ramp, etc.

The rest of the variables eliminated included the following. TRAFSGTP are used to describe traffic signal type at the site of a crash, such as intersection control, mid-block pedestrian control, emergency control or flashing beacon. In this study, all traffic signals are used for intersection control. TRAFCTR1 describes the first traffic control type at the site of a crash. In this study, this information is already known for each intersection. TRAFCHAR describes the geometric characteristics of roadway, such as straight or curve. Data analysis shows that there is very little variation for this variable. Straight situation covers 93.8% and 95.04% of crashes in the "before" and "after" data files respectively. NINSLEG records the number of intersection legs. The intersections of interest in the modeling process were analyzed as a whole, and not based on intersection types separately, such as three-leg intersections and four-leg intersections.

4.8.6. Variables Left in the Database

After the final screening, 19 variables were left in the database, including intersection identification variables, roadway characteristics variables, and crash counts variables. Table 4.11 lists these 19 variables. The final modeling database was built based on these variables.

Table 4.11. Variables List after Final Screening

No.	Variable	Description
1	ACCNUMB	Crash ID Number
2	DISTID	District Number
3	COUNTYID	County Number
4	SECID	Section Number
5	SUBSECID	Subsection Number
6	MILEPOST	Milepost of Crash Spot
7	ACCYEAR	Crash Year
8	ADT	ADT of Major Road
9	FEDHWY	Roadway Functional Level
10	URBRUR	Rural/urban
11	HARMEV1	First Harmful Event
12	SHOULDER	Shoulder Treatment Type
13	NUMLANES	Number of Lane on Major Road
14	DIVIDNOT	Presence of Median
15	LOCATYPE	Surrounding Land Use
16	POSTSPED	Posted Speed on Major Road
17	POINTIMP	Point of Impact
18	TURNONYR	Signal Activation Year
19	MILEPO	Milepost of Intersection Center

4.9. Converting Crash-based Database to Intersection-based Database

Once the final variable screening was completed, the next step was to convert the crash-based database to intersection-based database. In the intersection-based modeling database, a record corresponds to an intersection. Three types of variables were included in the modeling database: (1) intersection ID variables, (2) intersection characteristics variables, and (3) crash counts variables. The definitions of variables are shown in Table 4.12.

Table 4.12. Variables in the Final Modeling Database

Type	Variable	Description
Intersection ID Variables	DISTID	District Number
	COUNTYID	County Number
	SECID	Section Number
	SUBSECID	Subsection Number
	MILEPO	Milepost of Intersection
	TURNONYR	Signal Activation Year
Intersection Characteristics Variables	AVGADT	Average ADT of Major Road
	CLASS	Roadway Functional Level
	URBRUR	Rural/urban
	LOCATYPE	Surrounding Land Use
	LANE	Number of Lane on Major Road
	SPEED	Posted Speed on Major Road
	MEDIAN	Presence of Median
	SHOULDER	Shoulder Treatment Type
Crash Counts Variables	CRASH	Total Number of All Crashes
	AVGCRASH	Average Number of All Crashes Per Year
	ANGLE	Total Number of Angle Crashes
	AVGANG	Average Number of Angle Crashes per Year
	REAREND	Total Number of Rear-end Crashes
	AVGREAR	Average Number of Rear-end Crashes per Year
	LEFTTURN	Total Number of Left-turn Crashes
	AVGLEFT	Average Number of Left-turn Crashes per Year
	HEADON	Total Number of Head-on Crashes
	AVGHEAD	Average Number of Head-on Crashes per Year
	RIGHT	Total Number of Right-turn Crashes
	AVGRIGHT	Average Number of Right-turn Crashes per Year
	SIDESWIPE	Total Number of Sideswipe Crashes
	AVGSIDE	Average Number of Sideswipe Crashes per Year
	OTHER	Total Number of All Other Crashes
	AVGOTHER	Average Number of All Other Crashes per Year

The procedure to obtain the intersection-based database involved the selection of the three types of variables mentioned above for a three-year period for each intersection. Crash data for an intersection in any of the years considered could be zero, one or more crashes. This possibility of having different number of crashes also means that it could be zero, one or more crash records related to this intersection in the crash-based database. If there was zero crash or no crash at the intersection, the values of all the characteristic variables were set to missing. If there was one crash, the total number of crashes was equal to one. Crash type was identified based on the code of HARMEV1, which will be presented in the following paragraph. The values of the characteristic variables would be equal to the values of the corresponding variables in the crash record, e.g. NUMLANES corresponding to LANE. Finally, if there were more than one crash at the intersection, the total number of crashes and the total number of different types of crashes were determined base on HARMEV1. For the characteristic variables, inconsistency of the data among the crash records could be possible. This means that for each variable it could be one value for all the crash records, or as many values as the number of crash records. If all records had the same value, that value would be taken for the variable at this intersection. If the values are different, the value that appeared most frequently for a variable was chosen to represent that variable at the intersection.

The same method was applied to process the "before" data each one of the years considered. Once this task was completed, three "before" values were available for each variable. The next step was to calculate or select a value for each variable based on the three values. The values of the variables representing the average numbers of crashes per year (different types as well as all types), such as AVGCRASH, AVGANG, AVGLEFT, AVGREAR, AVGHEAD, AVGRIGHT, AVGSIDE and AVGOTHER, were calculated based on crash counts in each year. For example, there were two crashes in the first year, four in the second year, and five in the last year at this intersection; then the value of AVGCRASH was equal to 3.7. The value of AVGADT, representing average ADT on the major road, was calculated by averaging the three ADT values. If the value of ADT was missing in a year, then a reasonable value according to the ADT in a near year was assigned. For characteristic variables, one of the three values for each variable was selected because the three values were the same in most cases. If any inconsistency among the three values was found, the value with the highest frequency was chosen.

In regard to crash types, they were determined based on the variable HARMEV1. Coding system for this variable in the FDOT database was listed on the previous chapter on Table 3.2. The most important types of crashes that could occur at intersections include angle, left-turn, rear-end, head-on, right-turn, and sideswipe crashes. All of the other crash types were lumped together in category called "Other" crash type. In the modeling process, head-on, sideswipe and right-turn crashes were also added to “Other” crash type due to insufficient crash counts to be modeled separately. Summarily, the crash types selected for modeling include: all crashes, angle crashes, left-turn crashes, rear-end crashes, and other crashes.

Once the two steps were completed, the "before" value of the variables for this intersection were available. The "after" value of the variables for this intersection were determined in the same way except that all the values were based on a two-year period because only two-year "after" data were available in the second step. If the time frame was a one-year period, the values of the variables would be based on one-year data. This same procedure presented above was applied to other intersections in the intersection sample. Once this task was completed, the intersection-based modeling database was constructed.

4.10. Intersections with No Crash

For intersections without crash in the intersection influence area within the chosen time frame, FDOT database was analyzed again to find out the intersection characteristics. Two measures were used: either expanding the time frame or expanding the coverage of the intersection influence area.

To illustrate the procedure, an intersection with signals activated in 1996 is taken as an example. Assume there was no crash at this intersection from 1993 to 1995. In order to determine the intersection characteristics, the influence area was adjusted from 250 ft to a relatively longer distance, say, 300 ft or 400 ft, and the FDOT database was searched again. Usually, one or two crash would be found, and the intersection characteristics were identified based on these data. If this method failed, but there are some crash records after signalization, the intersection characteristics can be identified based on the crash data after signalization. The methods may have some drawbacks, but its effects are neglectable for modeling.

4.11. Closely Spaced Intersections

There are several pairs of intersections that are closely spaced. It is possible that the influence areas of the two intersections overlap. Additional work was done to make sure crashes were not counted twice. If so, the influence areas for those particular intersections were adjusted to avoid it happening. Among the closely spaced intersections in the intersection sample, some intersections were combined together as one intersection. For example, there are two intersections with the same district number, county number, section number and subsection number, the milepost of one intersection is 3.817, and the milepost of the other is 3.825. Thus, the center points of the two intersections are 42ft away. In data processing, they were combined as one intersection with the milepost being 3.821. Three pairs of intersections were found with this situation in the intersection sample.

4.12. Discarding Some Intersections

4.12.1. Intersections with Significant Geometric Improvements

The capability of the crash predictive models to explain the impacts of the signalization could be reduced if too many intersections with significant geometric improvement are kept in the intersection sample because the variation of crash frequency at intersections before and after the signalization could result from randomness, geometric changes, signalization and other factors. To avoid this, 14 intersections with significant geometric improvements in a three-year period before and/or after signalization were eliminated from the database.

Intersections with minor improvements from a safety point of view, such as installation of new signs, striping, resurfacing, extension of left-turn or right-turn lanes, etc., were not eliminated from the intersection sample. The extension of either left turn or right turn lanes were not taken into account because there was not information in regard to the length of the original turn lanes and the length of their extension, which does not allow to determine if the influence area considered for the research was exceeded. In regard to resurfacing, only two intersections had this improvement and a decision to keep them in the sample was made. On the other hand, intersections with significant geometric improvements such as reconstruction, change from 3-leg to 4-leg, adding left-turn lane, roadway widening, etc. were eliminated. Table 4.13 lists those intersections and their corresponding improvement.

4.12.2. Intersections on US-1 Bus-way Corridor

Seventeen intersections on US-1 bus-way corridor were eliminated from the intersection sample because there is no crash record in FDOT database for those intersections. Table 4.14 shows the ID variables of those intersections.

Table 4.13. Intersections with Significant Geometric Improvements

NO.	DISTID	COUNTYID	SECID	SUBSECID	MILEPO	TURNONYR	IMPROVEMENT
1	5	36	4	0	3473	90	Adding auxiliary lanes
2	6	87	1	0	8012	90	Adding additional left-turn lane
3	7	10	160	0	7156	92	Reconstruction
4	7	10	120	0	124	93	Recontruction
5	5	11	2	0	4514	94	Adding roadway connection
6	7	10	90	0	11389	94	Adding auxiliary lanes
7	7	10	90	0	11362	94	Adding auxiliary lanes
8	5	36	100	0	7956	95	Adding turn lanes
9	7	10	10	0	16953	95	Signal removed
10	7	14	90	0	16141	95	Adding additional connection
11	7	10	90	0	4368	96	Roadway widening
12	7	15	30	0	5794	96	Completely rebuilt
13	6	87	44	0	272	97	Adding connection
14	7	10	110	0	7424	96	Major improvement

Table 4.14. Intersections on US-1 Busway Corridor

NO.	DISTID	COUNTYID	SECID	SUBSECID	MILEPO	TURNONYR
1	6	87	207	0	369	97
2	6	87	207	0	1228	97
3	6	87	207	0	1734	97
4	6	87	207	0	1918	97
5	6	87	207	0	2385	97
6	6	87	207	0	2697	97
7	6	87	207	0	3066	97
8	6	87	207	0	3467	97
9	6	87	207	0	4189	97
10	6	87	207	0	4740	97
11	6	87	207	0	5311	97
12	6	87	207	0	5864	97
13	6	87	207	0	6148	97
14	6	87	207	0	6990	97
15	6	87	207	0	7548	97
16	6	87	207	0	7967	97
17	6	87	207	0	8360	97

4.12.3. Intersections without "Before" Data

There were eight intersections listed in Table 4.15 without "before" crash data. These intersections are located along the same roadway according to the ID variables. However, these intersections had relatively high crash experiences after signalization. Based on the findings, they were excluded from the intersection sample to avoid possible biased effects on the modeling.

Table 4.15. Intersections without "Before" Data

No.	DISTID	COUNTYID	SECID	SUBSECID	MILEPO	TURNONYR
1	1	16	6	0	215	91
2	1	16	6	0	596	91
3	1	16	6	0	1097	91
4	1	16	6	0	1223	91
5	1	16	6	0	1360	91
6	1	16	6	0	2484	91
7	1	16	6	0	3239	91
8	1	16	6	0	3740	91

4.13. Summary

Once the steps choosing time frames for crash analysis, identifying intersection related crashes, and selecting variables for the database were completed, the program SAS was chosen to conduct the database-building task. Data processing programs were written to retrieve data from the FDOT database, and generate two data files: the “before” and “after” crash data files. The "before" data file consists of 4565 crashes that occurred in the influence area of the 518 intersections within the "before" time frame, while the "after" data file consists of 6122 crashes within the “after” time frame.

After data reduction and analyses were completed, the final modeling database was built. The database consists of two intersection-based data files, one for the "before" period, and the other for the "after" period. In each data file, totally 447 intersections were included. For each intersection, 30 variables were used to record the safety-related information. The 30 variables are categorized into the following three groups: (1) intersection ID variables, including DISTID, COUNTYID, SECID, SUBSECID, MILEPO, and TURNONYR, (2) Intersection characteristic variables, including AVGADT, CLASS, URBRUR, LANE, LOCATYPE, SPEED, MEDIAN, and SHOULDER, and (3) Crash counts variables, including CRASH, AVGCRASH, ANGLE, AVGANG, REAREND, AVGREAR, LEFTTURN, AVGLEFT, HEADON, AVGHEAD, RIGHT, AVGRIGHT, SIDESWIPE, AVGSIDE, OTHER and AVGOTHER.

CHAPTER 5. CRASH DATA ANALYSIS

5.1. Analysis by Crash Type

Six crash types were selected for the before and after comparison. The six types are rear-end, right angle, left turn, right turn, sideswipe, and pedestrian. The mean values of annual number of crashes were calculated for the crash data of newly signalized intersections. Table 5.1 shows the results.

Table 5.1. Comparison of Average Number of Crashes for Different Crash Types

Period	Rear-end	Angle	Left turn	Right turn	Sideswipe	Pedestrian	Other
Before	0.95	1.05	0.96	0.12	0.18	0.07	0.25
After	1.93	0.91	0.80	0.10	0.25	0.067	0.65

Figure 5.1 presents the changes for each crash type for the intersections after signalization. Paired t-test was performed to verify whether the change was significant or not. Table 5.2 shows the results of the paired t-test. This table also presents the percent of change of the mean values, which is the absolute value between the "before" mean and "after" mean divided by the before mean value:

$$\text{Percent of Change} = |\,(\text{after mean} - \text{before mean})\,| \,/\, (\text{before mean}) \qquad (5\text{-}1)$$

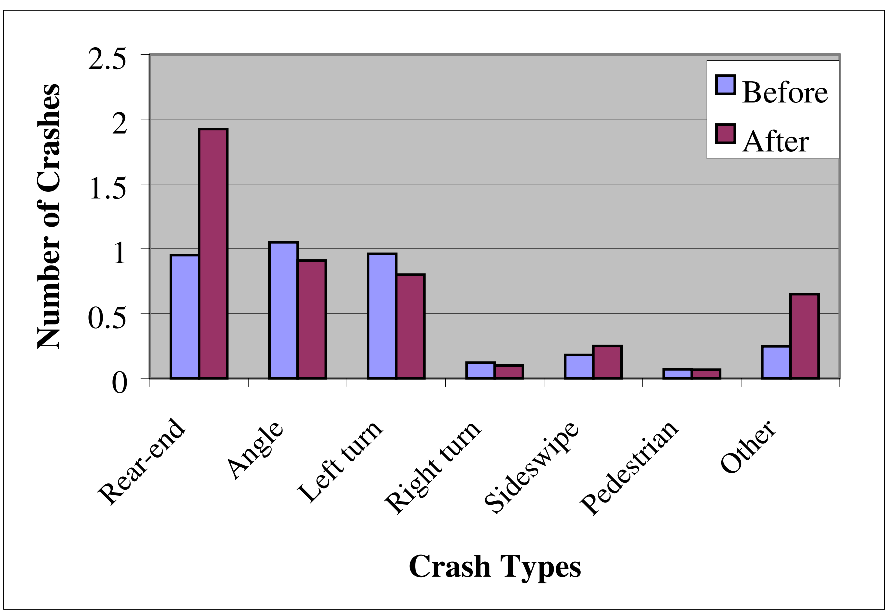

Figure 5.1. Before/After Average Number of Crashes for Different Crash Types

Table 5.2. Statistics for Average Number of Crashes by Crash Type

Crash Type	Before		After		T stat	T Critical two-tail	Significance of Difference	Increase or Decrease	Percent of Change
	Mean	Variance	Mean	Variance					
Rear-end	0.95	2.82	1.93	6.52	-7.18	1.96	Significant.	Increase	102%
Angle	1.05	2.19	0.91	1.32	1.99	1.96	Significant.	Decrease	14.30%
Left-turn	0.96	2.17	0.80	1.46	2.08	1.96	Significant.	Decrease	17%
Right turn	0.12	0.09	0.10	0.07	1.37	1.96	No Significant.	Decrease	18.6%
Sideswipe	018	0.13	0.25	0.23	-2.83	1.96	Significant.	Increase	41.9%
Pedestrian	0.07	0.078	0.07	0.04	0.208	1.96	No Significant.	Decrease	4.6%
All-other	0.25	0.45	0.65	0.63	-8.76	1.96	Significant.	Increase	163.3%

Results indicate that rear-end crashes would significantly increase after the signal is installed. The average number of rear-end crashes in the “after” period is twice as much as that in the “before” period. This increase in rear-end crashes due to signalization may be caused by the fact that vehicles on the major have to stop in the after period while in the before period no stopping was necessary. This increase already is of common knowledge to traffic engineers, but the 102% increase maybe higher than expected. Results also show that angle crashes, left turn crashes, right turn crashes, and crashes related with pedestrian decrease after signals are installed. The severity of some of these types of crashes may compensate the increase of the total number of crashes after signal installation. The increase of total number of crashes is caused basically by the increase of rear-end crashes.

Angle crashes are caused when one vehicle tries to cross the road, and a vehicle strikes it perpendicularly. After a signal is installed, angle crashes should decease because vehicles from the side street will be assigned the right of way so they could leave the side street and cross the main street without problems. From this study, it was found that the number of angle crashes decreased significantly with a 14% reduction. Left turn crashes were similar to angle crashes. It is known that after the signal is installed at the intersection, the opportunity of collision of left turn vehicles with vehicles at right-angle direction will decrease. But the possibility of collision of left turn vehicles with the vehicles coming in opposite direction depends on the signal operations (left turn treatment). If left turn movements are protected, left turn crashes will most likely decrease. In this study, left-turn crashes decreased significantly with a 17% reduction. In reference to the decrease in the number of right turn crashes, it could be a consequence of the right of way given by the signal to through movement and right turn from the same approach. This decrease of crashes is not statistically significant based on the paired t test. Sideswipe crashes occurred due to the lane changing near intersections. At signalized intersections, lane changing will be more difficult due to the stopping of other vehicles. The decrease of average number of crashes related with pedestrian may indicated that the traffic signal protects pedestrians. It was found that 4.6 percent of crashes related with pedestrian decreased after the signal was installed, but this decrease was not statistically significant.

In addition to the comparison of number of crashes, the average crash rates for the new signalized intersections were estimated based on the methodology explained earlier, and a before and after analysis was performed. Crash rates of rear-end crashes increased by 47.6%, which is lower than the 102% obtained for number of rear-end crashes. However, angle, left-turn, and right turn crash rates decreased in 29.02%, 37.6% and 50.2% respectively, which were more significant than their percentage decrease in number of crashes. Based on crash data, it is found that the change of sideswipe crash rates and crash rates related with pedestrian is not significant. The increase of crash rate of all other types of crashes is still very high with a 131.7% change. Figure 5.2 presents the changes of crash rates for different crash types from the before and after analysis. Table 5.3 presents the results of paired t-test for these changes.

5.2. Analysis by Crash Severity

For crash severity, the number of crashes and crash rates for the before and after periods were investigated to reveal the impacts of signalization in three categories: no injury, injury, and fatal. Here, injury crashes are the combination of possible injury, non-incapacitating injury, and incapacitating injury in the before-and-after analysis. Table 5.4 gives the changes of number of crashes and crash rates by different severities. The number of fatal crashes decreased by 13.2% and fatal crash rates decreased by 38% after the signal was installed. After signal installation, the number of no injury crashes increased by 30% and no injury crash rates by 14.8%. The number of injury crashes increased 17.2%, however, injury crash rates decreased by 5%. In conclusion, fatal crashes decreased, and no injury and injury crashes increased after signalization. Figures 5.3 and 5.4 show the results for number of crashes and crash rates by crash severity.

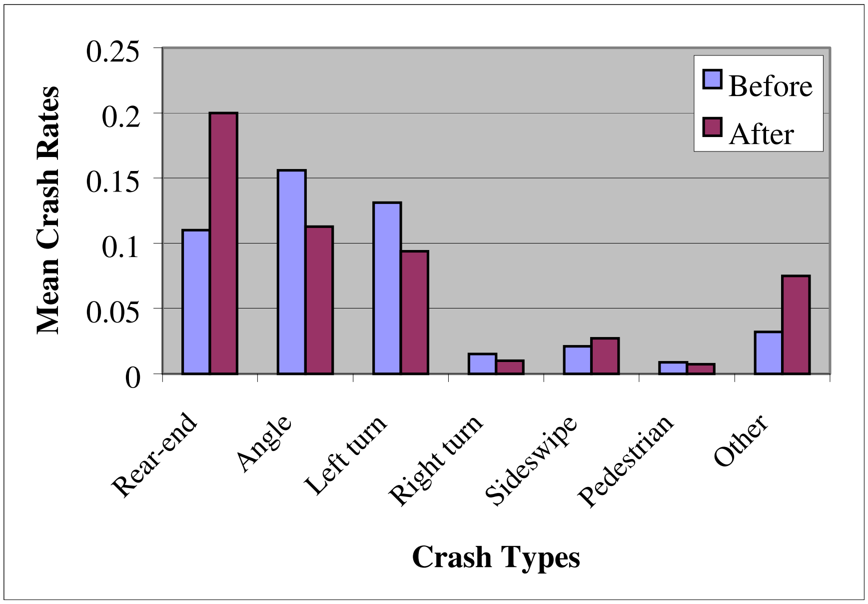

Figure 5.2. Before/After Average Crash Rates for Different Crash Types

Table 5.3. Statistics for Change of Average Crash Rates by Crash Type

Crash Type	Before		After		T stat	T Critical two-tail	Significance of Difference	Increase or Decrease	Percent of Change
	Mean	Variance	Mean	Variance					
Rear-end	0.11	0.12	0.20	0.04	-3.49	1.96	Significant.	Increase	47.6%
Angle	0.16	0.44	0.11	0.03	2.54	1.96	Significant.	Decrease	29.02%
Left-turn	0.13	0.15	0.10	0.02	3.25	1.96	Significant.	Decrease	37.6%
Right turn	0.02	0.00	0.01	0.00	3.45	1.96	Significant.	Decrease	50.2%
Sideswipe	0.02	0.01	0.03	0.00	-0.31	1.96	No Significant.	Increase	6.04%
Pedestrian	0.09	0.00	0.01	0.00	0.81	1.96	No Significant.	Decrease	17.4%
All-other	0.03	0.01	0.08	0.01	-6.82	1.96	Significant.	Increase	131.7%

Table 5.4. Statistics for Crashes and Crash Rates by Crash Severity

Crash Severity	Measure	Period	Observed Mean	Observed Variance	t-stat.	t Critical two-tail	Significance of Difference	Percent of Change
No Injury	Number of crashes	Before	1.38	4.23	-2.94	1.965	Significant Increase	30.0%
		After	1.79	5.97				
	Crash Rate	Before	0.18	0.067	-2.031	1.965	Significant Increase	14.8%
		After	0.19	0.061				
Injury	Number of crashes	Before	2.39	7.08	-2.51	1.965	Significant Increase	17.2%
		After	2.80	8.27				
	Crash Rate	Before	0.33	0.16	0.89	1.965	No Significant Decrease	5%
		After	0.32	0.08				
Fatal	Number of crashes	Before	0.06	0.03	0.67	1.965	No Significant Decrease	13.2%
		After	0.05	0.03				
	Crash Rate	Before	0.01	0.002	1.98	1.965	Significant Decrease	38%
		After	0.06	0.0004				

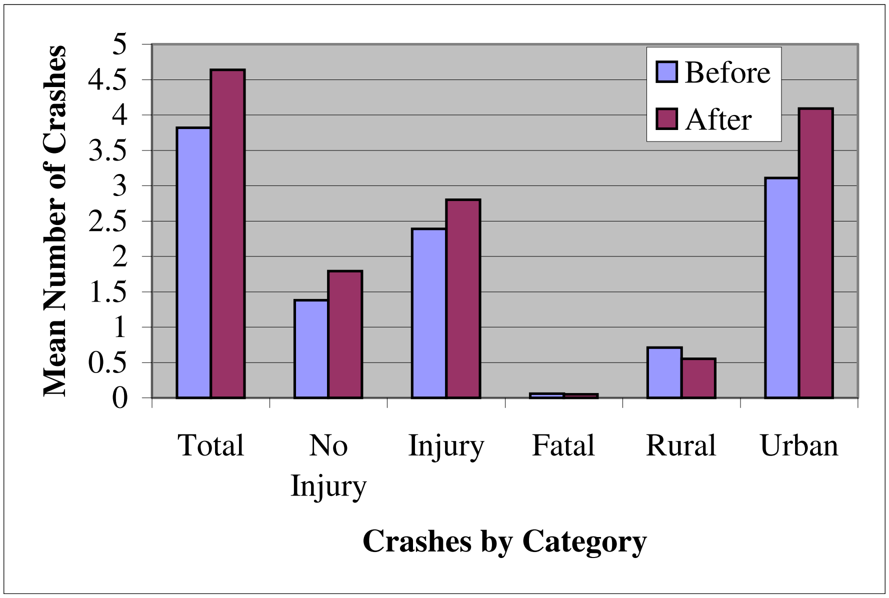

Figure 5.3. Before/After Average Number of Crashes Comparison

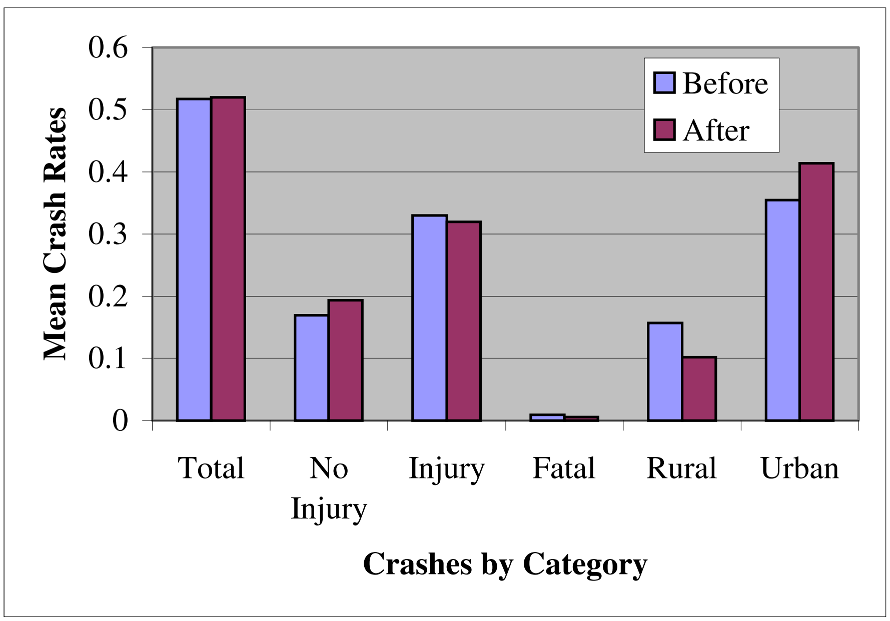

Figure 5.4. Before/After Average Crash Rates Comparison

5.3. Analysis for Surrounding Land Use

In general, urban areas tend to exhibit greater crash frequency than rural areas. Compared with rural areas, there are more driveways, and higher traffic volumes in urban areas with lower speeds. Based on the analysis of the crash data for surrounding land use, the impact of signalization on intersection crashes was found to be quite distinct between rural and urban areas. In rural areas, the number of crashes did not decrease significantly while the crash rates decreased significantly after signalization. In urban areas, the number of crashes and crash rates increased significantly. The results showed a decrease in the number of rural crashes, which was statistically insignificant, and a decrease in crash rates in rural areas, which was statistically significant. In regard to urban areas, the results analysis showed a statistically significant increase for number of crashes and crash rates. The reason for the increase of the number of crashes in urban areas could be related to the increase of specific types of crashes such as rear-end after signalization, which will have a greater increase in urban areas due to the high volume of vehicles. Table 5.5 presents the percent of change and results of paired t-test by number of crashes and crash rates for rural and urban areas. Figures 5.3 and 5.4 also present the results for number of crashes and crash rates by surrounding land use.

Table 5.5. Statistics for Crashes and Crash Rates for Surrounding Land Use

Crash location	Measure	Period	Observed Mean	Observed Variance	t-stat.	t Critical two-tail	Significance of Difference	Percent of Change
Rural	Number of crashes	Before	0.71	3.6	1.51	1.96	No Significant Decrease	30.00%
		After	0.55	2.48				
	Crash Rate	Before	0.16	0.235	2.302	1.965	Significant Decrease	14.80%
		After	0.10	0.085				
Urban	Number of crashes	Before	3.11	18.83	-3.43	1.96	Significant Increase	31.50%
		After	4.09	25.08				
	Crash Rate	Before	0.36	0.216	-2.227	1.965	Significant Increase	16.60%
		After	0.41	0.207				

5.4. Analysis for Total Number of Crashes and Crash Rates

Based on the mean value comparison for the before and after periods, total number of crashes significantly increased, while crash rates did not change. Table 5.6 lists the percent of change and the results of paired t-test by the total number of crashes and crash rates. These results are also show above on Figures 5.3 and 5.4.

Table 5.6. Statistics for Change of Total Number of Crashes and Crash Rates

Measure	Period	Observed Mean	Observed Variance	t-stat.	t Critical two-tail	Significance of Difference	Percent of Change
Total Number of crashes	Before	3.82	18.18	-3.71	1.96	Significant Increase	21%
	After	4.64	23.48				
Total Crash Rate	Before	0.517	0.348	-0.092	1.96	Significant Increase	0.58%
	After	0.520	0.219				

5.5. Crash Distributions for Before and After Conditions

Crash distribution modeling was performed for total number of crashes, crash types, crash severity and crash by surrounding land use. These distributions show how crashes varied and the type of distribution they followed. As examples, Figures 5.5 and 5.6 present the crash distributions for total number of crashes for before and after conditions. Both distributions show high number of intersections with low number of crashes. Similar results were founded for the rest of the crash distributions.

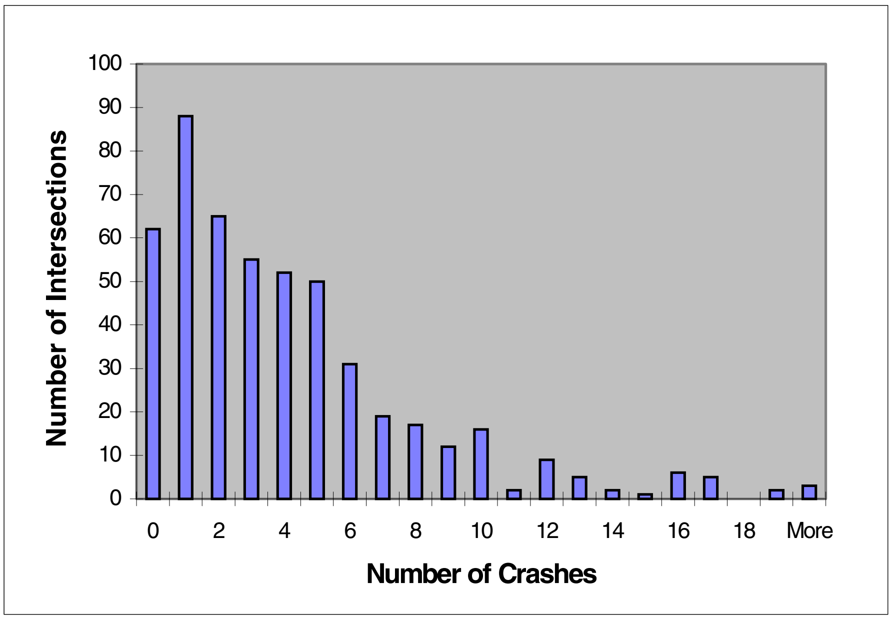

Figure 5.5. Total Number of Crashes Distribution Before Signalization

5.6. Distribution Fittings for Number of Crashes

Based on the frequency distributions and cumulative probability for total number of crashes, the mean and variance were calculated for the distribution fitting. The mean or expected value of the discrete random variable X, denoted as E(x), and the variance of x, denoted as V(x), are calculated as

$$E(x)= \sum_{x} x \times f(x) \quad (5\text{-}2)$$

$$V(x)= \sum_{x} (x - E(x))^2 \times f(x) \quad (5\text{-}3)$$

where,

f(x) = the probability of each random variable x.

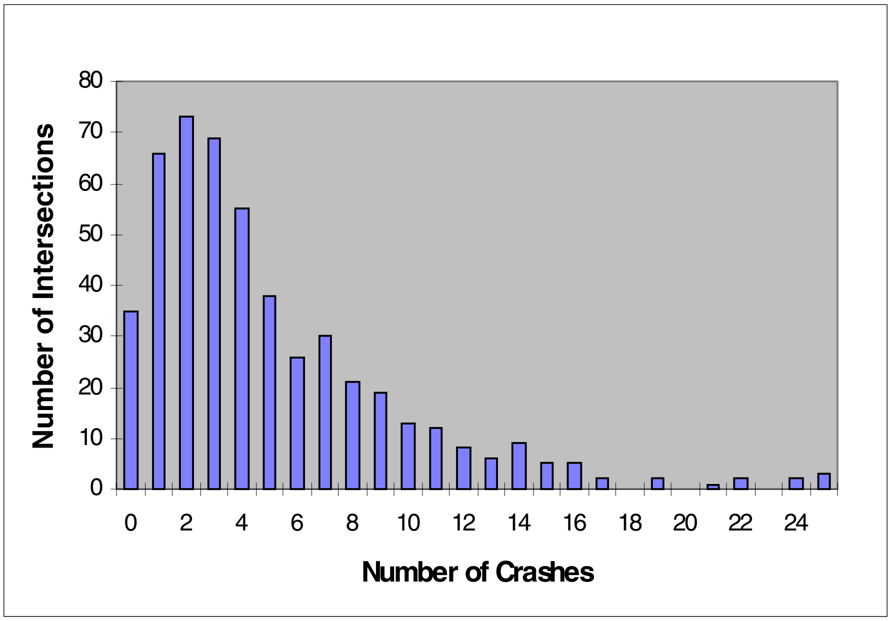

Figure 5.6. Total Number of Crashes Distribution After Signalization

Table 5.7 shows the procedure to get the mean and variance for total number of crashes for the before period.

Table 5.7. Mean and Variance of Total Number of Crashes for Before Period

X	Frequency	Cumulative %	F(x)	Xf(x)	$(x-E(x))^2$	$f(x)\ (x-mean)^2$
0	62	12.35%	12.35%	0.00	16.24	2.01
1	88	29.88%	17.53%	0.18	9.18	1.61
2	65	42.83%	12.95%	0.26	4.12	0.53
3	55	53.78%	10.96%	0.33	1.06	0.12
4	52	64.14%	10.36%	0.41	0.00	0.00
5	50	74.10%	9.96%	0.50	0.94	0.09
6	31	80.28%	6.18%	0.37	3.88	0.24
7	19	84.06%	3.78%	0.26	8.82	0.33
8	17	87.45%	3.39%	0.27	15.76	0.53
9	12	89.84%	2.39%	0.22	24.70	0.59
10	16	93.03%	3.19%	0.32	35.64	1.14
11	2	93.43%	0.40%	0.04	48.58	0.19
12	9	95.22%	1.79%	0.22	63.52	1.14
13	5	96.22%	1.00%	0.13	80.46	0.80
14	2	96.61%	0.40%	0.06	99.40	0.40
15	1	96.81%	0.20%	0.03	120.34	0.24
16	6	98.01%	1.20%	0.19	143.28	1.71
17	5	99.00%	1.00%	0.17	168.22	1.68
18	0	99.00%	0.00%	0.00	195.16	0.00
19	2	99.40%	0.40%	0.08	224.10	0.89
More	3	100.00%	0.60%			
	Total = 502		E (x) =	4.03	V(x)=14.24	

Using the observed mean and variance, the Poisson and Negative Binomial distributions were fitted to the crash data distribution for different crash types and severities. Table 5.8 demonstrates that the Chi-Square test for Poisson distribution fitted for total number of crashes for the before period. Thus,

$$f(x)\text{-Poisson} = \frac{e^{-\lambda}\lambda^{x}}{x!} = \frac{e^{-4.03}4.03^{x}}{x!}, \qquad x=0,\ 1,\ 2,\ \ldots\ldots, \tag{5-4}$$

where, λ is the mean value of observed data from Table 5.7, and λ = 4.03.

Table 5.8. Chi-Square Test for Poisson Distribution Fitted for Total Number of Crashes in Before Period

xr	F(i)	f(x)-Poisson	F(i)-f(x)	$(f(i)-f(x))^2$	$(f(i)-f(x))^2/f(x)$
0	12.35%	0.01777	0.10573	0.01118	0.62895
1	17.53%	0.07163	0.10367	0.01075	0.15004
2	12.95%	0.14434	-0.01485	0.00022	0.00153
3	10.96%	0.19389	-0.08433	0.00711	0.03668
4	10.36%	0.19534	-0.09176	0.00842	0.04310
5	9.96%	0.15745	-0.05785	0.00335	0.02125
6	6.18%	0.10575	-0.04400	0.00194	0.01831
7	3.78%	0.06088	-0.02303	0.00053	0.00871
8	3.39%	0.03067	0.00319	0.00001	0.00033
9	2.39%	0.01373	0.01017	0.00010	0.00753
10	3.19%	0.00553	0.02634	0.00069	0.12534
11	0.40%	0.00203	0.00196	0.00000	0.00189
12	1.79%	0.00068	0.01725	0.00030	0.43684
13	1.00%	0.00021	0.00975	0.00010	0.45024
14	0.40%	0.00006	0.00392	0.00002	0.25331
15	0.20%	0.00002	0.00198	0.00000	0.23910
16	1.20%	0.00000	0.01195	0.00014	34.71697
17	1.00%	0.00000	0.00996	0.00010	101.75062
18	0.00%	0.00000	0.00000	0.00000	0.00000
19	0.40%	0.00000	0.00398	0.00002	342.88459

The Chi-Square calculation value obtained from the Poisson distribution fitted for the observed total number of crashes was calculated with:

$$\chi_0^2 = \sum_{i=1}^{k} (\text{f(i)-f(x)})^2/\text{f(x)} \quad . \tag{5-5}$$

The value estimated of χ_0^2 is 481.77. This value is much bigger than the Chi-Square test value obtained from the Negative Binomial distribution fitting (0.131), as well as the Chi-

Square table value $\chi^2_{\alpha,k-p-1}$ =28.87 (α = 0.05, k = 20, p = 1). These Chi square results indicate that the hypothesis which state that the distribution of the total number of crashes in the before period is the hypothesized Poisson distribution is rejected.

Table 5.9 explains how the Chi-Square test is processed for the Negative Binomial distribution fitted for the total number of crashes for the intersections for the before period. As mentioned before, the Negative Binomial distribution has two parameters, mean E(x) and variance V(x). The probability function of X is:

$$f(x) = \binom{x-1}{r-1} p^r (1-p)^{x-r}, \qquad x = r, r+1, \ldots\ldots \tag{5-6}$$

Changing the scale in the previous equation by replacing x by x + r,

$$f(x) = \binom{x+r-1}{x} p^r (1-p)^{x}, \qquad x = 0, 1, 2, \ldots\ldots \tag{5-7}$$

In this case, from the observed mean E(x) and variance V(x), parameter *p* can be obtained from E(x)/V(x) and parameter r acquired from E(x)/(1/p-1) [21] [1] .

Table 5.9. Chi-Square Test for Negative Binomial Distribution
Fitted for Total Number of Crashes in Before Period

x	f(i)	f(x)-Negative Binomial.	f(i)-f(x)	$(f(i)-f(x))^2$	$(f(i)-f(x))^2/f(x)$
0	12.35%	0.07840	1	0.07840	0.04511
1	17.53%	0.05645	2	0.11290	0.06240
2	12.95%	0.04064	3	0.12193	0.00755
3	10.96%	0.02926	4	0.11705	-0.00749
4	10.36%	0.02107	5	0.10535	-0.00176
5	9.96%	0.01517	6	0.09102	0.00858
6	6.18%	0.01092	7	0.07646	-0.01470
7	3.78%	0.00786	8	0.06291	-0.02506
8	3.39%	0.00566	9	0.05096	-0.01709
9	2.39%	0.00408	10	0.04077	-0.01686
10	3.19%	0.00294	11	0.03229	-0.00041
11	.40%	0.00211	12	0.02536	-0.02138
12	1.79%	0.00152	13	0.01978	-0.00185
13	1.00%	0.00110	14	0.01534	-0.00538
14	.40%	0.00079	15	0.01183	-0.00785
15	.20%	0.00057	16	0.00909	-0.00710
16	1.20%	0.00041	17	0.00695	0.00500
17	1.00%	0.00029	18	0.00530	0.00466
18	.00%	0.00021	19	0.00403	-0.00403
19	.40%	0.00015	20	0.00305	0.00093

In Table 5.9, p = 4.03/14.24 = 0.28, r = E(x)/(1/p-1)= 2.

$$f(x)\text{ -Negative Binomial} = \frac{(2+x-1)!}{(2-1)!x!}(0.28)^2(1-0.28)^x$$

Chi-Square calculation value estimated from the Negative Binomial distribution fitted with observed total number of crashes is χ_0^2 = 0.131. This value is smaller than the Chi-Square table value $\chi^2_{\alpha,k-p-1}$ = 27.59 (α = 0.05, k = 20, p=2), which indicates that the hypothesis that the distribution of the total number of crashes for the before period is the hypothesized Negative Binomial distribution will not be rejected. Finally, it could be concluded that the Negative Binomial distribution is better to fit the distribution of total number of crashes at the intersections from the Chi-Square test comparison. The same result was obtained for the total number of crashes in the after period, with Chi-Square calculation value 45842.89 for Poisson and 0.063 for Negative Binomial, respectively. Figures 5.7 and 5.8 present the graphs of frequency distributions, which illustrate the same outcome for distribution fitting of the total number of crashes. Table 5.10 exhibits the Chi-Square test for fitting Poisson and Negative Binomial distributions by different crash types, crash severities, and crash by surrounding land use. If both Poisson and Negative Binomial distribution were not rejected by the Chi-square test ($\chi_0^2 < \chi^2_{\alpha,k-p-1}$), the distribution with smaller Chi-square calculation value was selected as the fitted distribution. Table 5.10 shows that the Negative Binomial distributions are selected as the fitted distribution to fit the number of crash distributions.

5.7. Fitting for Crash Rate Distributions

The method used for the number of crashes distribution fitting was also applied for crash rate distribution fittings. As examples for frequency distributions, crash rate distributions for total crashes are presented for the before and after periods in Figures 5.9 and 5.10, respectively.

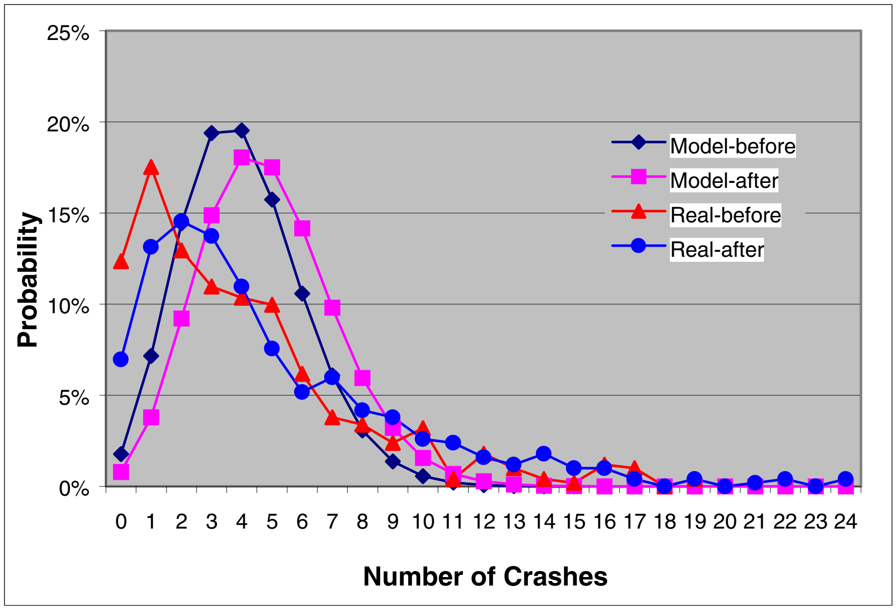

Figure 5.7. Comparison of Poisson Distribution and
Observed Total Number of Crashes Distribution

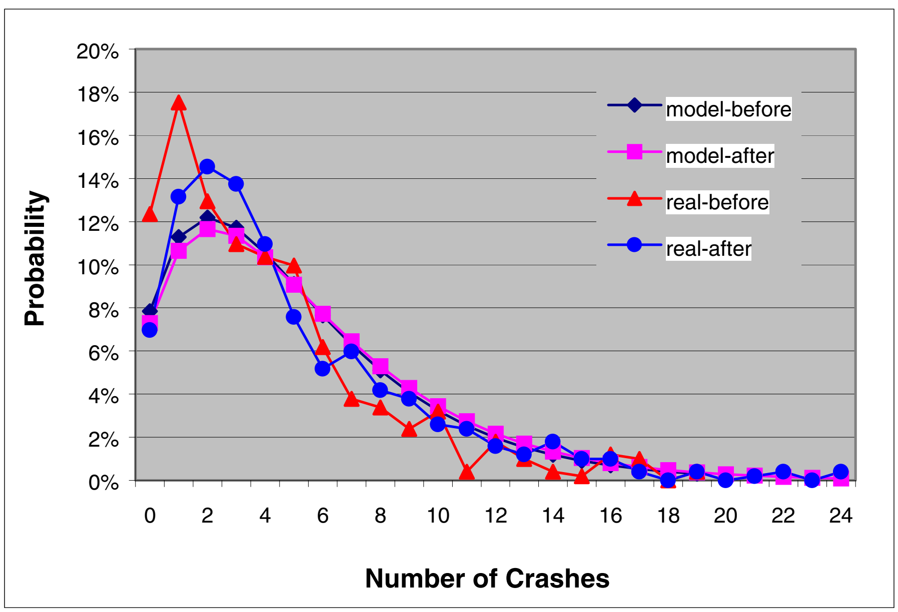

Figure 5.8. Comparison of Negative Binomial Distribution and
Observed Total Number of Crashes Distribution

Table 5.10. Chi-Square Test Comparison for Poisson and Negative Binomial Distribution Fitting for Number of Crashes Distributions

Category	Period	Poisson: Chi-Square Calculation. χ_0^2	Poisson: Chi-Square Table Value $\chi^2_{\alpha,k-p-1}$	Negative Binomial: Chi-Square Calculation. χ_0^2	Negative Binomial: Chi-Square Table Value $\chi^2_{\alpha,k-p-1}$	Distribution Selected
Total Number of crashes	Before	481.77	28.87	0.131	27.59	Negative Binomial
	After	45842.89	35.17	0.063	33.92	Negative Binomial.
Rear End	Before	146.29	18.31	0.208	16.92	Negative Binomial
	After	6.32	18.31	0.075	16.92	Negative Binomial
Angle	Before	0.3	12.59	0.033	11.07	Negative Binomial
	After	0.09	12.59	0.055	11.07	Negative Binomial
Left Turn	Before	0.294	12.59	0.012	11.07	Negative Binomial
	After	0.116	12.59	0.058	11.07	Negative Binomial
Other	Before	0.186	12.59	0.053	11.07	Negative Binomial
	After	0.176	12.59	0.148	11.07	Negative Binomial
No Injury	Before	398.5	21.03	0.077	19.68	Negative Binomial
	After	26.24	21.03	0.26	19.68	Negative Binomial
Injury	Before	1.39	18.31	0.087	16.92	Negative Binomial
	After	169.3	26.3	0.074	25	Negative Binomial
Rural	Before	0.03	9.488	0.004	7.815	Negative Binomial
	After	0.229	9.488	0.012	7.815	Negative Binomial
Urban	Before	0.3	14.07	0.01	12.59	Negative Binomial
	After	0.254	14.07	0.026	12.59	Negative Binomial

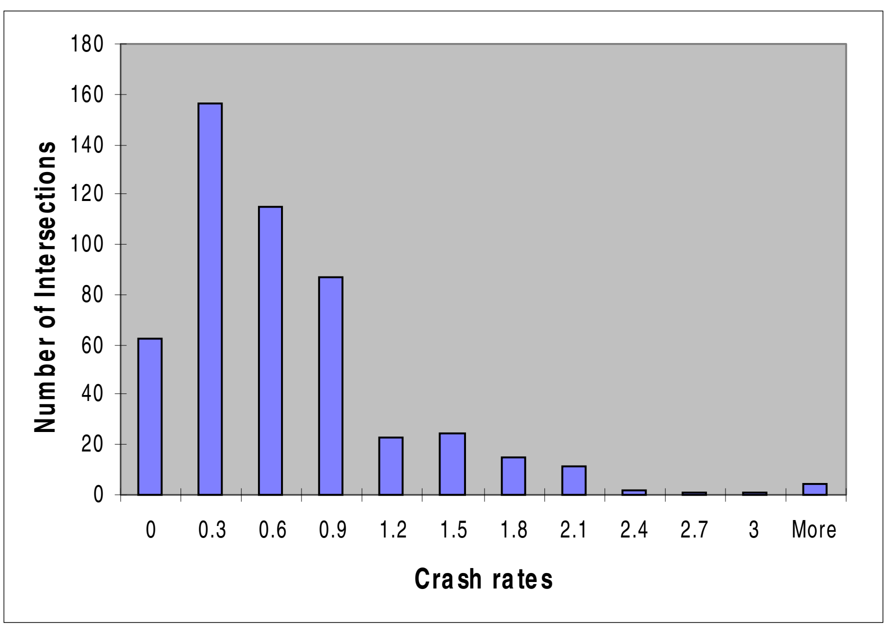

Figure 5.9. Total Crash Rate Distribution Before Signalization

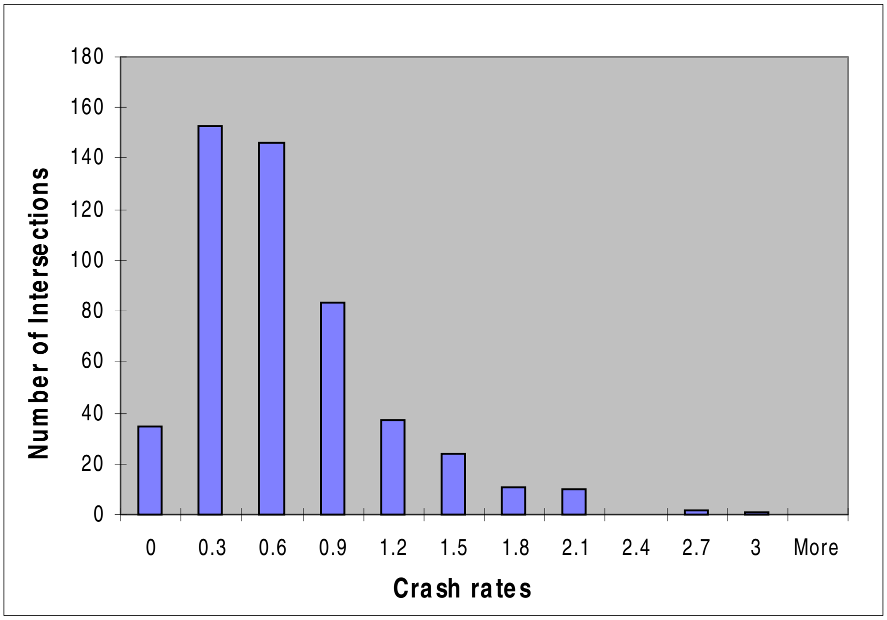

Figure 5.10. Total Crash Rate Distribution After Signalization

Similar to the results for total number of crashes, the Negative Binomial distribution was found to be the best fitting distribution for most of the cases for crash rate distribution fitting, except for urban crash rate distribution. In previous research studies, the ratio of variance and mean were used to choose the fitted Negative Binomial and Poisson distribution. The selection criteria are:

1. Poisson distribution, if variance/mean ratio =1.0, and
2. Negative Binomial distribution, if variance/mean ratio >1.0

In this study, Negative Binomial was found to be better than Poisson distribution even with the variance/mean ratio closed to 1.0, due to the tail part. But when variance/mean ratio is less than 1, the Negative Binomial distribution cannot be employed because the parameter r is less than 0. Thus, Poisson distribution was used to fit these cases. Poisson distribution was used for the crash rate distribution in urban area because the variance/mean ratio was less than 1. Figures 5.11 and 5.12 illustrate the same results. Table 5.11 lists the Chi-Square test values for Poisson and Negative Binomial Distributions. Eight of them were fitted to Negative Binomial distributions, and the last one was fitted to a Poisson distribution.

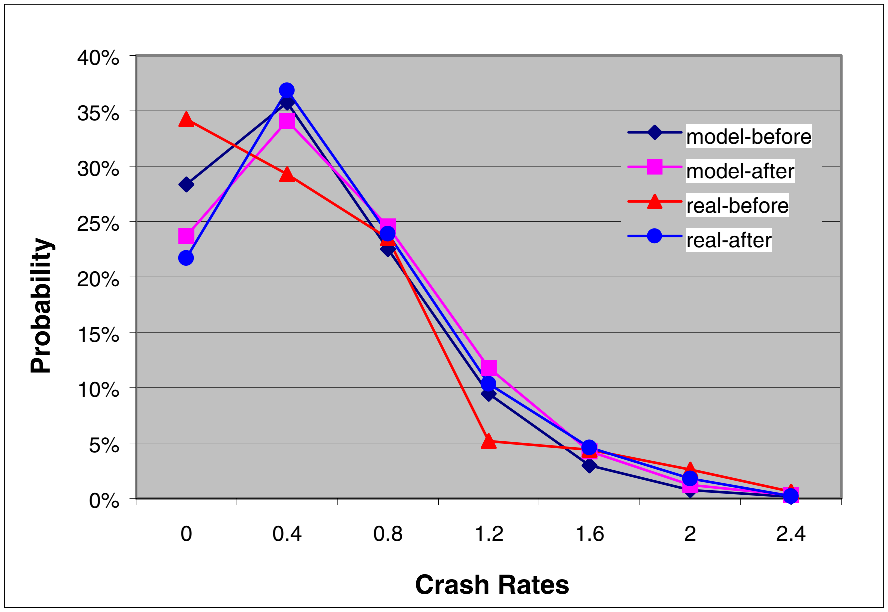

Figure 5.11. Comparison of Poisson Probability Distribution and Observed Urban Crash Rate Probability Distribution

Figure 5.12. Comparison of Poisson Cumulative Distribution and Observed Urban Crash Rate Cumulative Distribution

Table 5.11. Chi-Square Test Comparison for Poisson and Negative Binomial Distribution Fitting for Crash Rate Distributions

Category	Period	Poisson: Chi-Square Calculation. χ_0^2	Poisson: Chi-Square Table Value $\chi^2_{\alpha,k-p-1}$	Negative Binomial: Chi-Square Calculation. χ_0^2	Negative Binomial: Chi-Square Table Value $\chi^2_{\alpha,k-p-1}$	Distribution Selected
Total Crashes	Before	0.21	16.92	0.069	15.51	Negative Binomial
	After	0.156	16.92	0.093	15.51	Negative Binomial
Rear End	Before	0.058	11.07	0.0085	9.488	Negative Binomial
	After	0.021	11.07	0.012	9.488	Negative Binomial
Angle	Before	12.36	16.92	0.095	15.51	Negative Binomial
	After	11.5	16.92	0.164	15.51	Negative Binomial
Left Turn	Before	1.87	15.51	0.064	14.07	Negative Binomial
	After	2.03	15.51	0.256	14.07	Negative Binomial
Other	Before	1.372	15.51	0.019	14.07	Negative Binomial
	After	0.05	15.51	0.022	14.07	Negative Binomial
No Injury	Before	2.01	16.92	0.087	15.51	Negative Binomial
	After	0.834	16.92	0.098	15.51	Negative Binomial
Injury	Before	1.40	16.92	0.263	15.51	Negative Binomial
	After	0.084	16.92	0.072	15.51	Negative Binomial
Rural	Before	0.676	14.07	0.0025	12.59	Negative Binomial
	After	0.373	11.07	0.101	9.488	Negative Binomial
Urban	Before	0.109	11.07	\	9.488	Poisson
	After	0.009	11.07	\	9.488	Poisson

5.8. The 50th and 85th Percentile Values of Crashes

The 50th and 85th percentile values of each crash distribution were obtained for the before and after periods. The 85th percentile is the point where 85 percent of the crashes at an intersection will occur either at or below this measurement. This value is often used in engineering analysis because the data in the top 15 percent, considered the top portion of the population, is not targeted in design. Table 5.12 and 5.13 present the parameters used in the distribution equations, and the 50th and 85th percentile values for number of crashes and crash rate, respectively

Table 5.14 shows the percentage change for each 50th and 85th percentile value. These percentile values for the number of crashes and crash rates for before and after periods are shown from Figure 5.13 to 5.16. In the fitted Negative Binomial distributions, the increase of the total number of crashes after the installation of signals is not considerable. Based on the Negative Binomial distributions, the number of crashes and crash rates for rear-end crash type increased significantly, and left turn crashes deceased in a big extent However, the 50th percentile of the number of angle crashes is found to increase in the observed crash distribution and the fitted Negative Binomial distribution. Parameters r and p are calculated from observed the mean and variance of the distribution.

Table 5.12. Equation and 50th and 85th Percentile Values for Number of Crashes Distribution

Category	Before				After			
	Equation		50%	85%	Equation		50%	85%
	r	p			r	p		
Total Number of crashes	2	0.28	4.3	8.6	2	0.27	4.5	8.9
Rear End	1	0.55	0.6	1.6	2	0.49	1.2	3.7
Angle	2	0.62	0.4	2.2	6	0.84	0.5	1.8
Left Turn	2	0.64	0.5	2.0	6	0.86	0.4	1.6
Other	5	0.82	0.7	1.9	15	0.92	1.0	2.0
No Injury	5	0.66	0.4	2.7	2	0.45	1.7	4.8
Injury	1	0.40	1.6	2.8	2	0.42	1.3	4.8
Rural	9	0.88	1.2	4.0	2	0.64	0.7	3.9
Urban	3	0.64	2.8	8.4	4	0.69	3.0	8.6

Table 5.13. Equation and 50th and 85th Percentile Values for Crash Rate Distribution

Category	Before				After			
	Equation		50th	85th	Equation		50th	85th
	r	p			r	p		
Total Crash Rate	5	0.71	0.39	1.01	15	0.87	0.46	0.10
Rear End	4	0.78	0.06	0.28	20	0.92	0.15	0.40
Angle	2	0.52	0.15	0.48	4	0.76	0.07	0.30
Left Turn	1	0.44	0.03	0.20	1	0.54	0.00	0.15
Other	2	0.59	0.05	0.02	10	0.87	0.07	0.02
No Injury	2	0.58	0.05	0.25	2	0.53	0.09	0.30
Injury	8	0.79	0.27	0.675	24	0.92	0.27	0.66
Rural	2	0.51	0.20	0.77	2	0.56	0.14	0.60
Urban	λ=1.26		0.25	0.25	λ=0.75		0.30	0.90

Table 5.14. Before and After Difference for 50th and 85th Percentile Values

Type	Percentile	Total	Rear End	Angle	Left Turn	Other	No Injury	Injury	Rural	Urban
Number of crashes	50th	5%	100%	13%	-20%	43%	325%	-19%	-42%	7%
	85th	3%	131%	-18%	-20%	5%	78%	71%	-3%	2%
Crash Rate	50th	18%	150%	-53%	-100%	40%	70%	0%	-30%	20%
	85th	2%	45%	-38%	-25%	0%	20%	-2%	-22%	20%

Similar to the mean value comparison results, the number of no injury crashes and crash rates increased significantly while injury crash rates did not increase for the 50th percentile value and decreased for the 85th percentile value. Thus, signalization would improve safety at the intersections in rural areas. However, the number of crashes and crash rates would increase at intersections in urban areas after signal installation.

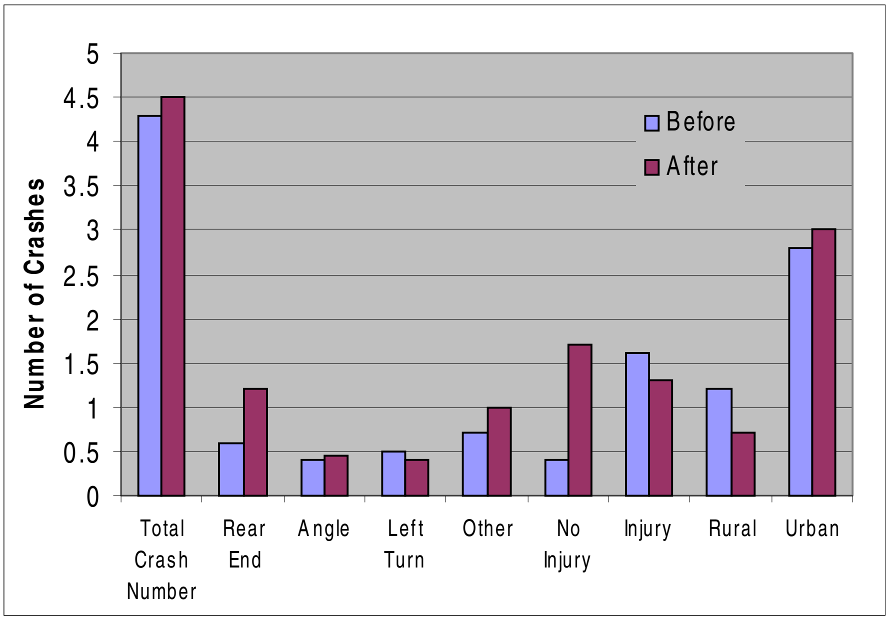

Figure 5.13. 50th Percentile Before/After Comparison for Number of Crashes

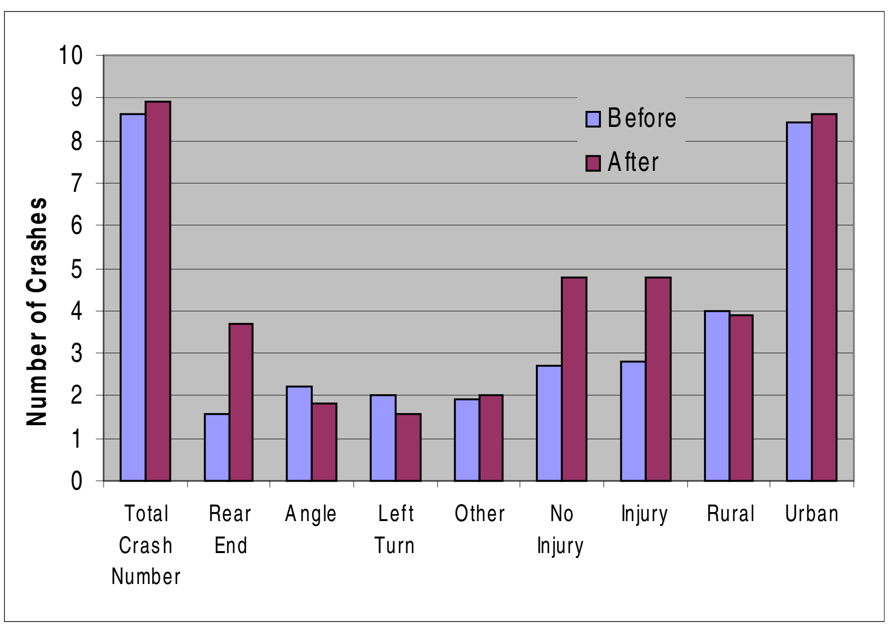

Figure 5.14. 85th Percentile Before/After Comparison for Number of Crashes

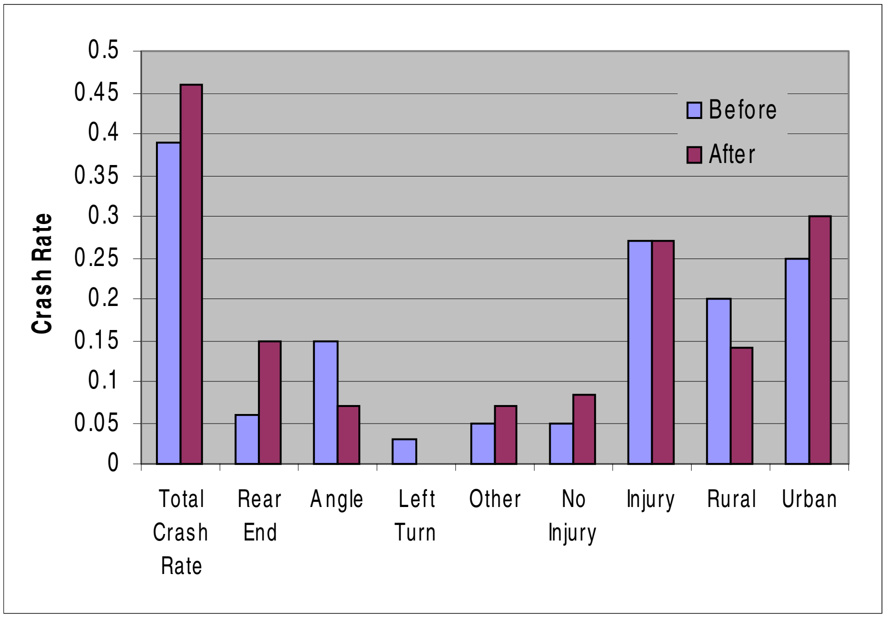

Figure 5.15. 50th Percentile Before/After Comparison for Crash Rate

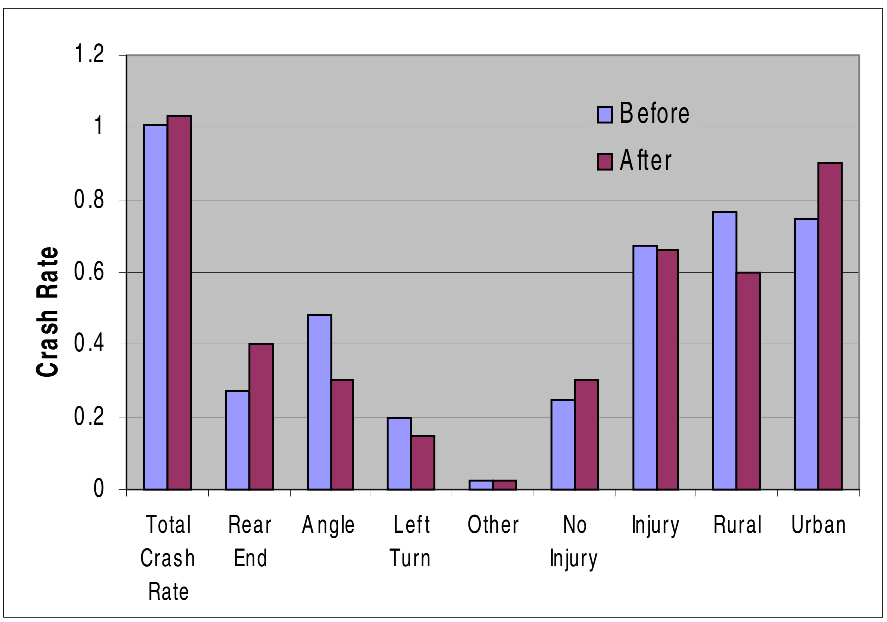

Figure 5.16. 85th Percentile Before/After Comparison for Crash Rate

CHAPTER 6. RESULTS FOR STATISTICAL MODELS

6.1. Dependent Variables

The dependent variables adopted in the modeling process included: (1) average number of all crashes per year before and after signalization, (2) average number of angle crashes per year before and after signalization, (3) average number of left-turn crashes per year before and after signalization, (4) average number of rear-end crashes per year before and after signalization, and (5) average number of all other crashes per year before and after signalization. It is important to note that "all other" crashes include all crashes except for angle, left-turn and rear-end crashes. The determination of the dependent variables was based on data analyses, which is presented in the following paragraphs.

Originally, the most common crash types such as angle, left-turn, rear-end, head-on, right-turn, and sideswipe were analyzed to develop predictive models, separately. However, based on data analysis, it was found that head-on, right-turn and sideswipe crash counts were not statistically sufficient. Out of the 447 intersections, about 90% of the intersections had no head-on crash, the remaining 10% of the intersections had more than zero but no more than one average head-on crash per year. About 80% of the intersections had no right-turn crash, the remaining 20% of the intersections had more than zero but no more than one average crash per year. More than 60% of the intersections had no sideswipe crashes, and about 30% of the intersections had more than zero but no more than one average sideswipe crash per year. Therefore, lack of variation as well as insufficient crash counts for the three types of crashes made it impossible to do modeling, separately. Only angle, left-turn, and rear-end crashes were chosen to develop predictive models, separately. Based on data analyses, the remaining types of crashes were aggregated in one category called the "all other" crash type. Thus, totally five types of crashes were chosen to be dependent variables to develop the predictive models.

The summary descriptive statistics for the dependent variables are shown in Table 6.1. It can be found that the mean value of the average number of angle and left-turn crashes per year after signalization was lower than the mean value before signalization. However, the mean value of the average number of rear-end crashes per year after signalization was significantly larger than the mean value before signalization. Overall, the mean value of the average number of all crashes per year after signalization was larger than the mean value before signalization. These results are in line with the results shown in the previous chapter and the common knowledge of traffic engineers.

It was found that the maximum value of the average number of crashes was much higher than the median value, which implies that there are some "black spots" in the intersection sample. The abnormal higher values may be a result of particular reasons different to normal traffic conditions at intersections. In the modeling process, the "black spots" were taken out from the intersection sample.

6.2. Predictor Variables

The selection of predictor variables was based on the available data and engineering judgment. The task was carried out through the database building process for the modeling. After the modeling database was built, totally eight variables were available. The following subsections provide detailed description and analyses for the eight variables selected.

6.2.1. ADT on the Major Road

Traffic volume is the most significant factor contributing to crash occurrence. The change of traffic volume entering an intersection imposes multiple effects on the traffic operations and safety at the intersection. In this project ADT data on the major road were available at each intersection, and it was represented by the variable AVGADT. The descriptive statistics of AVGADT are provided in Table 6.2.

ADT value was transformed from continuous to discrete value because the results of developed models will be tabulated for application by traffic engineers so that traffic engineers could easily apply the level of traffic volume (low, medium or high) to the models rather than search for the accurate ADT volume. Therefore, the variable AVGADT was converted to discrete variable based on its distribution. The 25% quantile was about 14,000 vpd, and the 75% quantile, 32,000 vpd. Thus, the thresholds to divide the ADT into low, medium, and high categories were set to 15,000 vpd and 30,000 vpd. Table 6.3 shows the range of each level and the value used in the models.

Table 6.1. Descriptive Statistics for Dependent Variables

Descriptive Statistics	Average Number of Crashes per Year									
	All Crashes		Angle		Left-turn		Rear-end		All Other	
	Before	After	Before	After	Before	After	Before	After	Before	After
Mean	4.075	4.937	1.111	0.94	1.045	0.856	0.99	2.068	0.929	1.072
Standard Deviation	4.255	4.957	1.478	1.138	1.525	1.257	1.679	2.655	1.152	1.216
Number of Intersections	447	447	447	447	447	447	447	447	447	447
Minimum	0	0	0	0	0	0	0	0	0	0
Maximum	34	41.667	8.668	7.333	10	12	18.667	28.667	7.333	9.333
Median	3	3.333	0.5	0.667	0.5	0.333	0.667	1.333	0.667	0.667
Skewness	2.465	2.514	1.889	1.99	2.689	3.478	5.12	3.865	2.306	2.272
Kurtosis	10.162	10.644	3.77	4.98	9.392	19.329	40.874	27.102	7.078	7.488

Table 6.2. Descriptive Statistics for the Variable AVGADT

Descriptive Statistics		"Before" Data	"After" Data
Number of Intersections		447	447
Mean (vpd)		24,304	26,878
Standard Deviation		14,924	15,713
Quantiles (vpd)	100% (Maximum)	116,266	125,519
	99%	70,056	74,423
	95%	54,042	55,146
	90%	44,517	48,606
	75%	31,467	34,167
	50% (Median)	21,833	23,833
	25%	13,613	14,501
	10%	8,599	9,703
	5%	6,961	7,500
	1%	2,160	5,300
	0% (Minimum)	927	2,167

Table 6.3 Levels of the Variable ADT in Models

Level	Traffic Volume Range	Value Used in Models
Low	< 15,000 vpd	0
Medium	15,000 vpd ~ 30,000 vpd	1
High	≥ 30,000 vph	2

6.2.2. Surrounding Land Use

Surrounding land use refers to urban or rural area. Generally, more crashes were observed in urban areas. Furthermore, in urban areas traffic volumes are higher, there are more lanes on the roadway, driveways are spaced more closely within the influence area of an intersection, turning traffic volumes are higher due to the complex travel destinations, and so on. On the other hand, vehicle speeds in urban areas usually are lower, congestion situations are more common, and proportion of heavy vehicles is lower. Thus, the effect of surrounding land use on intersection crash frequency is very complicated. The best way to explore the possible answer is to perform data analyses.

The variable URBRUR in the database describes whether an intersection is located in urban area or in rural area. Table 6.4 provides the descriptive statistics for this variable.

Roughly, the intersections were evenly split both before and after signalization. In the Table, frequency before and after refers to the number of intersections within each specific area, urban and rural for the before and after signalization period.

Table 6.4. Descriptive Statistics for the Variable URBRUR

Type	Value	Frequency		Percentage	
		Before	After	Before	After
Rural	0	244	243	54.59%	54.36%
Urban	1	203	204	45.41%	45.64%

6.2.3. Location Type

The variable LOCATYPE describes the land use of the area near the intersection including business, residential, shopping and recreational. As an environmental factor, land use can capture the effects of the area and its impacts on driver driving behavior. In residential areas, a relatively higher percentage of drivers are familiar with the operations of an intersection. Also, the access to the roadway tends to be regulated effectively within the influence area of an intersection. In business areas, to some extent turning movements could be performed more often, and there are more access points within the influence area of an intersection, that usually leads to more conflicts on the roadway.

Table 6.5 presents the descriptive statistics of the variable LOCATYPE, and Table 6.6 shows the levels and values used in modeling. The conversion from three levels to two levels was based on preliminary modeling efforts showing that the best results were obtained by this treatment. In the Table, frequency before and after refers to the number of intersections within each specific type, business or other for the before and after signalization period.

Table 6.5. Descriptive Statistics for the Variable LOCATYPE

Type	Value	Frequency		Percentage	
		Before	After	Before	After
Business	1	347	356	77.63%	79.64%
Residential	2	49	46	10.96%	10.29%
Other	3	51	45	11.41%	10.07%

Table 6.6. Levels of the Variable LOCATYPE in Models

Level	Value	Frequency		Percentage	
		Before	After	Before	After
Business	1	347	356	77.63%	79.64%
Other	2	100	91	22.37%	20.36%

6.2.4. Number of Lanes on the Major Road

The number of lanes on the major road is one of the most important geometric factors in explaining crash occurrence. Although there is correlation, which was found not to be high, between the number of lanes and traffic volume, it is not always true that higher volume imply more lanes. Together with traffic volume, it determines the level of service at an intersection, which directly influences and relates to the intersection safety. Level of service may have been an interesting variable to include in the models but this information was not available in the database.

Table 6.7 presents the descriptive statistics for the variable LANE that describes the number of lanes on the major road in both directions. Two-lane, four-lane, and six-lane are the most common cases. Table 6.8 shows the levels of this variable used in the modeling process. Based on preliminary modeling efforts, in order to get the best modeling results four lanes were used as threshold to divide the data into two categories. In the Table, frequency before and after refers to the number of intersections within each specific number of lanes, > 4 or ≤ 4 lanes for the before and after signalization period.

Table 6.7. Descriptive Statistics for the Variable LANE

Number of Lanes on Major Road	Frequency		Percentage	
	Before	After	Before	After
2	112	83	25.06%	18.57%
3	12	16	2.68%	3.58%
4	257	257	57.49%	57.49%
5	1	8	0.22%	1.79%
6	58	75	12.98%	16.78%
7	0	0	0.00%	0.00%
8	7	8	1.57%	1.79%

Table 6.8. Levels of the Variable LANE in Models

Level	Value	Frequency		Percentage	
		Before	After	Before	After
> 4 lanes	1	66	91	14.77%	20.36%
<= 4 lanes	0	381	356	85.23%	79.64%

6.2.5. Posted Speed on the Major Road

Posted speed is an important traffic speed control factor for traffic safety analysis. Usually, it is believed that crashes are more likely to occur at higher speed, which actually is not well documented. However, common engineering knowledge is that high speed more likely results in severe crashes. Also, the effect of posted speed on crash occurrences is more significant at unsignalized intersections than at signalized intersections. From another point of view, drivers tend to travel at speeds in which they feel comfortable given the prevailing conditions. Therefore, lower posted speed more likely promotes speed differential that is generally more closely associated with crashes. Table 6.9 presents the descriptive statistics for the variable SPEED, which describes the posted speed on the major road. Table 6.10 shows the levels of SPEED used in modeling. In the Table, frequency before and after refers to the number of intersections within each specific category, > 45mph or ≤ 45mph for the before and after signalization period.

Table 6.9. Descriptive Statistics for the Variable SPEED

Posted Speed on the Major Road	Frequency		Percentage	
	Before	After	Before	After
15	1	0	0.22%	0.00%
25	8	5	1.79%	1.12%
30	37	18	8.28%	4.03%
35	60	59	13.42%	13.20%
40	57	59	12.75%	13.20%
45	182	200	40.72%	44.74%
50	20	25	4.47%	5.59%

6.2.6. Presence of Median on Major Road

Another important roadway geometric factor to be considered is whether the roadway is divided or not. Generally, roadways having more lanes to carry higher traffic volume are divided and probably located in urban areas. On the contrary, roadways having fewer lanes and serving low traffic volume are undivided and probably located in rural areas. For the effect of median on intersection safety, there is no clear answer to whether the presence of median on the major road would increase or decrease the crash frequency. It also depends on many other factors such as traffic control, crash type, and so on. Table 6.11 gives the descriptive statistics of the variable MEDIAN and corresponding values for modeling. In the Table, frequency before and after refers to the number of intersections within each specific level, divided or undivided for the before and after signalization period.

Table 6.10. Levels of the Variable SPEED in Models

Level	Value	Frequency		Percentage	
		Before	After	Before	After
> 45 mph	1	102	106	22.82%	23.71%
≤ 45 mph	0	345	341	77.18%	76.29%

Table 6.11. Descriptive Statistics for the Variable MEDIAN

Level	Value	Frequency		Percentage	
		Before	After	Before	After
Divided	1	282	283	63.09%	63.31%
Undivided	0	165	164	36.91%	36.69%

6.2.7. Shoulder Treatment

Shoulder treatment is another interesting factor to include during modeling. Paved shoulder makes the drivers traveling on the right lane to feel safer. In some cases, paved shoulder can provide space to accommodate right-turn vehicles at intersections or vehicles traveling out of the proper lane. Unpaved shoulder could worsen the consequence once a vehicle runs off the pavement, especially for inexperienced drivers. Table 6.12 presents the descriptive statistics of the variable SHOULDER, indicating the type of shoulder treatment at an intersection. Curbed shoulders were combined with unpaved shoulders considering the fact that curbed shoulder also provides restraint on the lateral movement of vehicles traveling on the right lane, which could have similar effects on safety. Table 6.13 shows the levels of the variable SHOULDER used in modeling. These levels already included the combination of the unpaved and curb shoulder into the other category. In the Table, frequency before and after refers to the number of intersections within each specific type, paved and unpaved for the before and after signalization period.

Table 6.12. Descriptive Statistics for the Variable SHOULDER

Type	Value	Frequency		Percentage	
		Before	After	Before	After
Paved	1	134	148	29.98%	33.11%
Unpaved	2	292	210	65.32%	46.98%
Curb	3	21	89	4.70%	19.91%

Table 6.13. Levels of the Variable SHOULDER in Models

Level	Value	Frequency		Percentage	
		Before	After	Before	After
Paved	1	134	148	29.98%	33.11%
Other	2	313	299	70.02%	66.89%

6.2.8. Functional Class of Major Road

The variable describing the functional class of the major road is also included in the modeling database. According to this variable, the major road of an intersection can be identified as either arterial or collector. Originally, this variable was treated as one of the predictor variables. However, preliminary modeling results showed that in any case its effect was insignificant. Also, the classification of the functional class of a roadway usually is determined based on its functional position within the roadway network rather than roadway geometric characteristics. With these limitations, this variable was excluded from the predictor variables.

6.3. Crash Frequency Distributions

Prior to the statistical modeling, the general shapes of crash frequency distributions were assessed in order to provide the basis for crash distribution assumptions for modeling.

Figures 6.1 through 6.5 show the statistical results for all crashes, angle crashes, left-turn crashes, rear-end crashes, and all other crashes.

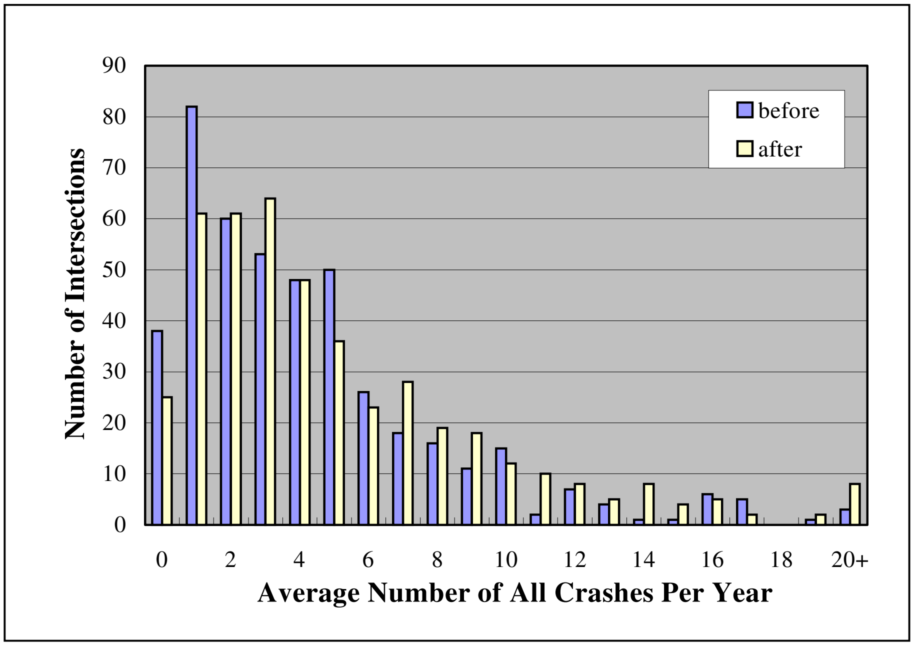

Figure 6.1. Frequency Distribution for All Crashes

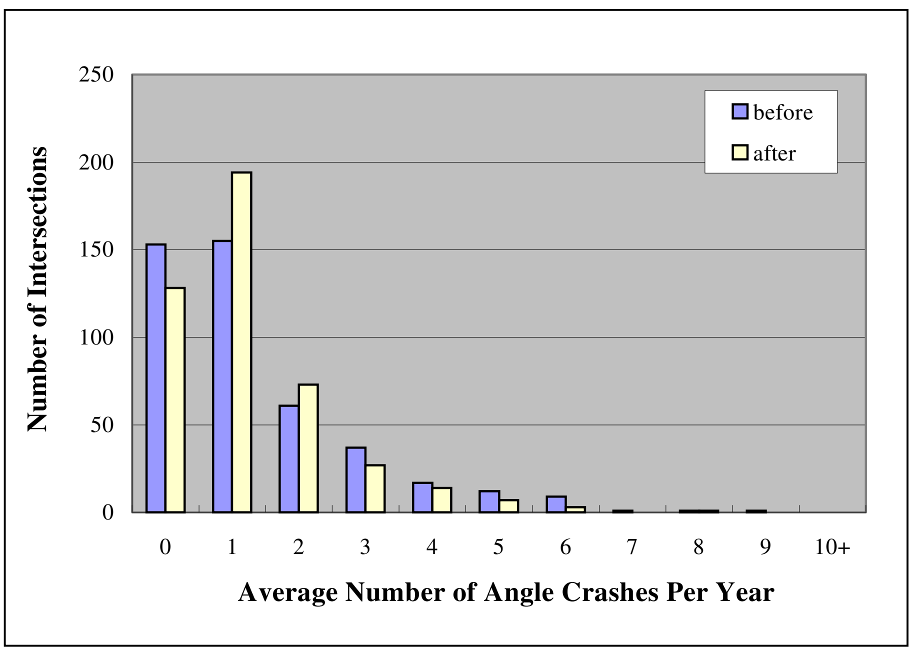

Figure 6.2. Frequency Distribution for Angle Crashes

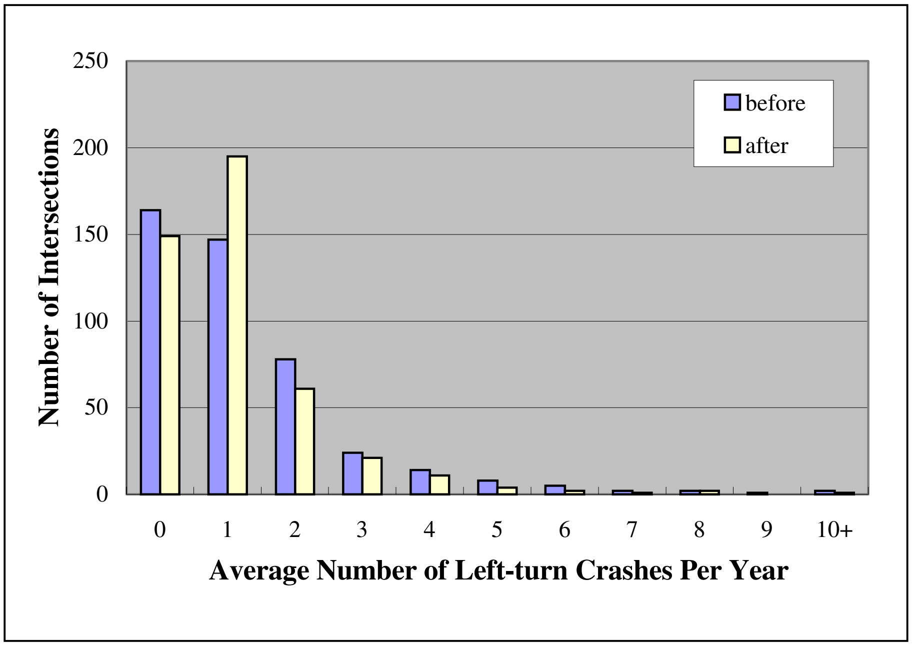

Figure 6.3. Frequency Distribution for Left-turn Crashes

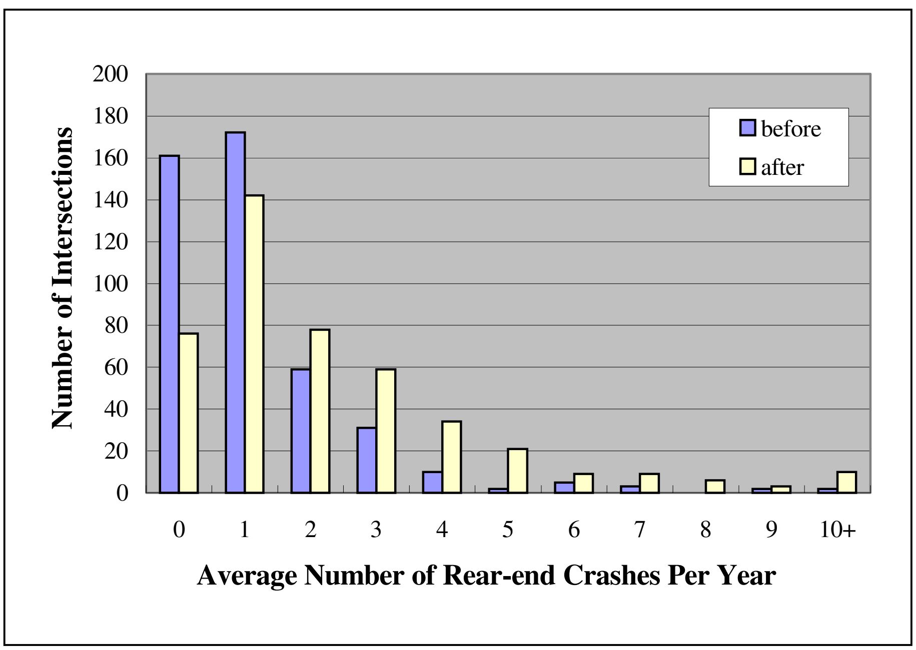

Figure 6.4. Frequency Distribution for Rear-end Crashes

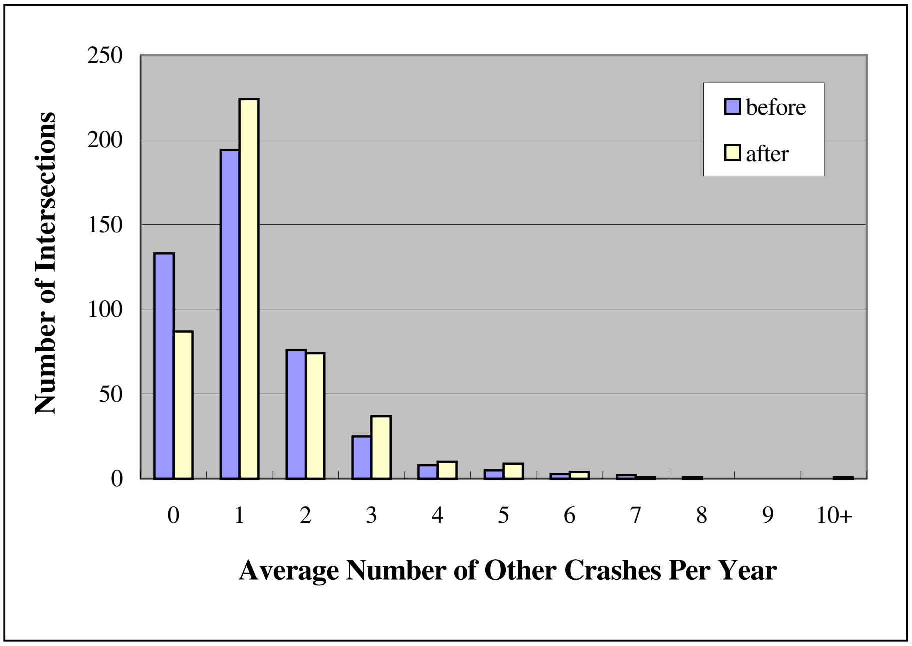

Figure 6.5. Frequency Distribution for Other Crashes

In Figure 6.1, it is clearly shown that a large number of intersections had no or low crash experience, and the distributions for the before and after period seems to follow the Poisson distribution. This was also statistically proven in phase one. Also, it can be found visually that the number of signalized intersections with high crash experience were more than the number of unsignalized intersections, which is consistent with the statistics shown in Table 6.1. The distributions shown in Figures 6.2 through 6.5 also follow the Poisson distribution. In Figure 6.2, the number of intersections with more than 3 angle crashes decreases after signalization while the number of intersections with 1 or 2 angle crashes increases, which leads to the decrease of the overall number of angle crashes. A similar situation can be found for left-turn crashes shown in Figure 6.3. However, the trend goes to the opposite direction for rear-end crashes as presented in Figure 6.4. The number of intersections with no or one rear-end crash decreases after signalization while the number of intersections with more than two rear-end crashes increases significantly, which leads to the increase of the overall number of rear-end crashes. In Figure 6.5, the change pattern for all other crashes is not very obvious, only showing slight increase of the number of intersections with more than one crash after signalization.

6.4. Crash Predictive Modeling

The following section presents the detailed modeling process for all crashes: For angle crashes, left-turn crashes, rear-end crashes, and all other crashes, the rest of the sections present a summary of the modeling. The parameter estimations were performed with the GENMOD procedure of the SAS statistical software package. Mean deviance and Pearson's Chi-square ratio were adopted as the criteria to test over-dispersion of crash data. Backward elimination method was used through the regression process to remove statistically insignificant predictor variables. Deviance, Pearson's Chi-square, Pearson's R-square and likelihood ratio index were adopted to evaluate the goodness-of-fit of developed models.

6.4.1. Models for All Crashes

6.4.1.1. The Model for Before Signalization

First, all seven predictor variables were included in the regression equation. The Poisson regression was performed as the initial step. Initial Poisson regression results provided the basis to test whether the crash data were over-dispersed. Two statistics were adopted as the criteria to assess the over-dispersion: mean deviance and Pearson's Chi-square ratio. Generally, the mean deviance and the Pearson's Chi-square ratio should be close to one or within the range between 0.8 and 1.2 in order to consider the Poisson model appropriate to fit the data. If the mean deviance ant the Pearson’s Chi-square ratio values exceed one, the data are considered to display extra variation or over-dispersion relative to the Poisson model. If the values are less than one, the data are said to display under-dispersion relative to the Poisson model. In Table 6.14, the mean deviance for the initial Poisson model is 3.209, and the Pearson's Chi-square ratio is 3.524, which indicate that the extra variation exists in the "before" data. An initial negative binomial regression was performed as an alternative to Poisson model; and the mean deviance and Pearson's Chi-square ratio were calculated again. As shown in Table 6.14, the mean deviance and Pearson's Chi-square ratio for the negative binomial model are very close to one, which indicate that the negative binomial model was an appropriate choice.

Based on the results of the negative binomial regression, Pearson residual for each observation was calculated and plotted in Figure 6.6. Pearson residual was used to check the model fit, where the best-fitted data points should be around the zero line in the plot. A 20 percent significance level was assumed to keep the parameters estimated in the model. The choice of a 20 percent significance level or 80 percent confidence level would allow to include several predictor variables in the predictive model that may improve the overall predictive ability of the model than if a more restrictive significance level was considered. The results of the negative binomial regression are presented in Table 6.15. The explanations of the contents of Table 6.15 are listed on Table 6.16.

Table 6.14. Criteria For Assessing Over-Dispersion (All, Before)

Criterion	DOF	Poisson Model (Initial)		NEGATIVE BINOMIAL Model (Initial)	
		Value	Value/DOF	Value	Value/DOF
Deviance	439	1408.571	3.209	481.505	1.097
Pearson's Chi-square	439	1546.937	3.524	466.218	1.062

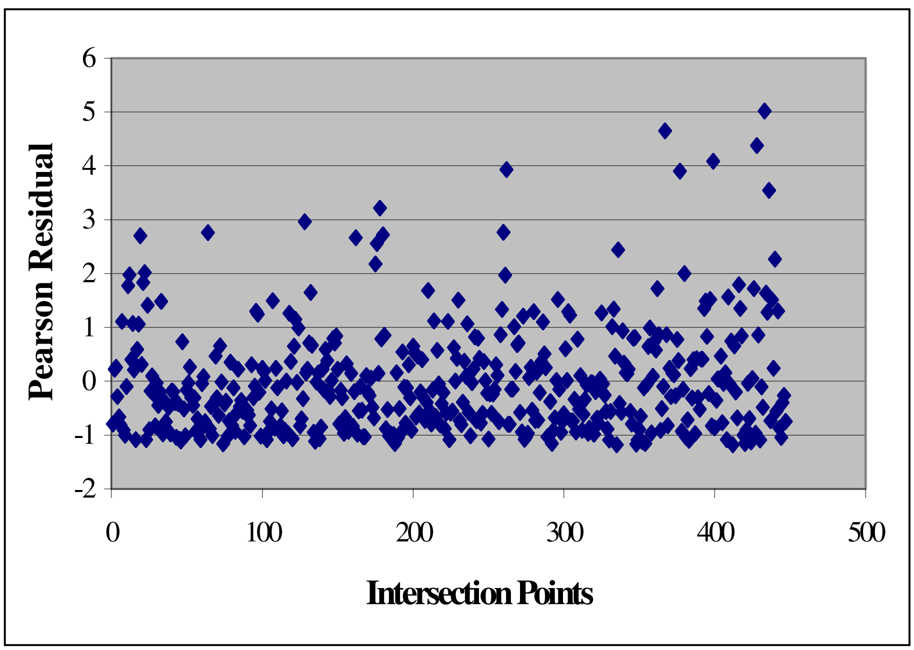

Figure 6.6. Pearson Residual for Initial Negative Binomial Model (All, Before)

Table 6.15. Estimated Parameters of the Negatie Binomial Model (All, Before)

Predictor Variable	Name	Level	Value	DOF	Coefficent Estimate	Standard Error	Relative Effect	Chi-Square	Pr>Chisq
Intercept				1	0.6827	0.1229		30.86	<0.0001
Average ADT	AVGADT	<15,000 vpd	0	1	0.2777	0.0641	1.32	18.74	<0.0001
		15,000~30,000	1						
		>=30,000 vpd	2						
Surrounding Land Use	URBRUR	Urban	1	1	0.1193	0.0866	1.13	1.90	0.1683
		Rural	0						
Location Type	LOCATYPE	Business	1	1	0.1705	0.1082	1.19	2.48	0.1151
		Other	0						
Number of Lanes on Major Road	LANE	> 4	1	1	0.2614	0.1214	1.30	4.63	0.0313
		<= 4	0						
Posted Speed on Major Road	SPEED	> 45	1	1	-0.1695	0.1071	0.84	2.51	0.1133
		<=45	0						
Presence of Median on Major Road	MEDIAN	Divided	1	1	0.2752	0.0954	1.32	8.32	0.0039
		Undivided	0						
Shoulder Treatment	SHOULD	Paved	1	1	-0.1679	0.0954	0.85	3.10	0.0785
		Other	0						
Dispersion Parameter				1	0.4480	0.0522			

Table 6.16 Explanation of Contents of the Results

Column	*Explanation*
Predictor Variable	Describe the variables associated with estimated parameters. The INTERCEPT represents the intercept in the regression equation. The over-dispersion parameter is an additional parameter estimated in the negative binomial model relative to Poisson model.
Name	Variable name in the modeling database.
Level	Represents the levels of the variable. Please note that ADT was treated as discrete variable with three values: 0, 1 and 2. Other variables were treated as categorical variables with two levels.
Value	Value of variable to be input in the developed model.
DOF	Degrees of freedom associated with each parameter estimate. Each categorical variable has k-1 degrees of freedom; k represents the levels of the variable. The intercept has one degree of freedom.
Parameter Estimate	Estimated parameters.
Standard Error	Estimated standard deviation associated with each parameter.
Relative Effect	Exponent of the estimated parameter of the variable. Represents the effects of different levels.
Chi-square	Chi-square test statistic for testing that the parameter is 0. This was computed as the square of the ratio of the parameter estimate divided by its standard error.
Pr > Chi-Sq	The probability of obtaining a Chi-square statistic greater than that observed given that the true parameter is 0. A small p-value is evidence for concluding that the parameter is not 0.

The following step is to assess the goodness-of-fit of the model. Four statistics, including deviance, Pearson's Chi-square, Pearson's R-square, and likelihood ratio index, were adopted. Table 6.17 presents the four statistics for the "before" negative binomial model of all crashes. Both the mean deviance and Pearson's Chi-square ratio are close to one, and the Pearson's R- square and the likelihood ratio index are around 20%. The statistics indicate that the developed model has satisfactory capability in fitting the "before" data and explaining the variation of the data.

In addition to statistical justification, the model should also satisfy engineering judgment. This can be assessed by examining the relative effect of each variable. For example, the relative effect of ADT is 1.32, which means that the average number of crashes would increase by 32% if the ADT increase from low to medium level, given all other variables constant. Similarly, if ADT increases from medium to high level, the average number of crashes would increase by 32%. An intersection in urban area would have 13% more crashes than the similar intersection in rural area. Intersections located in business area are more likely to have crashes by 19% than intersections located in other area, e.g. residential area. The same trend applies to other factors such as number of lanes. However, intersections with posted speed more than 45 mph would have 16% fewer crashes than similar intersections with posted speed less than or equal to 45 mph. Even though this result may not be as expected, the analysis of the data indicated this trend, where intersections with high-posted speed experienced lower number of crashes than intersections with low-posted speed. This could be explained by the fact that in areas with higher volume, the posted speed is lower than in areas with lower traffic volume. Higher volume usually means high number of crashes. It is also interesting to find that intersections with median on the major road would have 32% more crashes than without median. The presence of median is generally an indicator that intersections are located in areas with high traffic volume, probably with several lanes, which will produce more conflicts and thus more crashes. Finally, intersections with paved shoulder would have 15% fewer crashes than the intersections with other types of shoulder.

Table 6.17. Criteria for Assessing the Goodness-of-Fit
(All, Before, Negative Binomial Model)

Item	Value
Number of Observations (n)	439
Number of Predictor Variables in Model	7
Number of Parameters in Model (p)	7
Degree of Freedom (n-p-1)	431
Log-likelihood at Convergence	-914.001
Restricted Log-likelihood	-1149.14
Deviance	458.399
Deviance/(n-p-1)	1.064
Pearson Chi-square	413.318
Pearson Chi-square/(n-p-1)	0.959
Pearson R-square	20.33%
Likelihood Ratio Index	20.46%

6.4.1.2. The Model for After Signalization

The model for “after” signalization for all crashes was developed using a similar procedure as the model “before” signalization. Firstly, all predictor variables were included and Poisson regression was performed. Based on the results of the initial Poisson model, the mean deviance and Pearson's Chi-square ratio were calculated. Table 6.18 shows the results. Both statistics were larger than 3.0 showing that the "after" data were over-dispersed. As an alternative, negative binomial (NB) regression was performed, and the mean deviance and Pearson's Chi-square ratio were calculated again based on the negative binomial model. Both statistics are also presented in Table 6.18, these values were close to one indicating that the negative binomial regression is an appropriate choice. Then, Pearson residuals were calculated for each observation based on the results of negative binomial regression and plotted in Figure 6.7. Most of the points clustered between the -1 and 1 lines, indicating that the model fits the data satisfactorily

Table 6.18. Criteria For Assessing Over-Dispersion (All, After)

Criterion	DOF	Poisson Model (Initial)		Negative Binomial Model (Initial)	
		Value	Value/DOF	Value	Value/DOF
Deviance	439	1378.279	3.140	470.613	1.072
Pearson's Chi-square	439	1509.108	3.438	477.831	1.088

Of the seven-predictor variables in the model, the estimated parameter for posted speed is 0.004, with Chi-square statistic equal to 0.0000 and p-value equal to 0.9665, which means that the effect of posted speed is extremely insignificant. The negative binomial regression was run again after the predictor variable posted speed was removed from the regression equation. The results are shown in Table 6.19. After the comparison of the two sets of results, it was found that removing posted speed from the model had very few effects on other variables. Among the estimated parameters presented in Table 6.19, the parameter of the variable urban/rural is significant at 22 percent confidence level. Considering that it is only slightly lower than the adopted 20 percent confidence level, the variable has been kept in the model. The parameters for ADT, number of lanes on major road, and presence of median are significant at 5% significance level. The parameter of LOCATYPE is significant at 10% significance level. The parameter of shoulder treatment is significant at 20% significance level. The dispersion parameter is 0.3450 showing that the “after” data is over-dispersed relative to Poisson distribution.

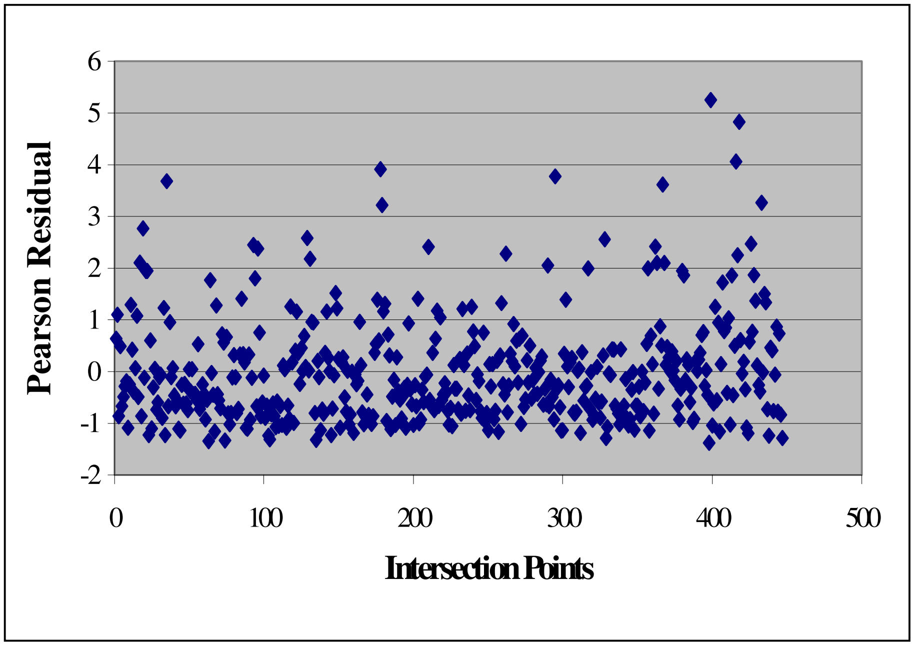

Figure 6.7. Pearson Residual for Initial Negative Binomial Model (All, After)

Also, the goodness-of-fit statistics were calculated for the final negative binomial model and are presented in Table 6.20. The mean deviance and Pearson's Chi-square ratio are close to one, which indicates that the negative binomial model fits the "after" data very well. The Pearson R-square value is equal to 34.84% and the likelihood ratio index is equal to 27.75 %, representing that the model has a satisfactory ability in explaining the variation of the "after" data. Each variable in the final " after" model shows the similar relative effect on crash occurrence at intersections as the final "before" model.

Table 6.19. Estimated Parameters of the Negative Binomial Model (All, After)

Predictor Variable	Name	Level	Value	DOF	Coefficent Estimate	Standard Error	Relative Effect	Chi-Square	Pr>Chisq
Intercept				1	0.5718	0.1142		25.08	<0.0001
Average ADT	AVGADT	<15,000 vpd	0	1	0.4868	0.0578	1.63	71.04	<0.0001
		15,000~30,000	1						
		>=30,000 vpd	2						
Surrounding Land Use	URBRUR	Urban	1	1	0.0949	0.0766	1.10	1.53	0.2158
		Rural	0						
Location Type	LOCATYPE	Business	1	1	0.1728	0.0966	1.19	3.20	0.0737
		Other	0						
Number of Lanes on Major Road	LANE	> 4	1	1	0.2654	0.0928	1.30	8.18	0.0042
		<= 4	0						
Posted Speed on Major Road	SPEED	> 45	1	1	~	~	~	~	~
		<=45	0						
Presence of Median on Major Road	MEDIAN	Divided	1	1	0.1845	0.0876	1.20	4.44	0.0351
		Undivided	0						
Shoulder Treatment	SHOULD	Paved	1	1	-0.1102	0.0813	0.90	1.84	0.1754
		Other	0						
Dispersion Parameter				1	0.3450	0.0400			

Table 6.20. Criteria for Assessing the Goodness-of-fit
(All, After, Negative Binomial Model)

Item	Value
Number of Observations (n)	438
Number of Predictor Variables in Model	6
Number of Parameters in Model (p)	6
Degree of Freedom (n-p-1)	431
Log-likelihood at Convergence	-911.514
Restricted Log-likelihood	-1261.53
Deviance	439.624
Deviance/(n-p-1)	1.020
Pearson Chi-square	420.532
Pearson Chi-square/(n-p-1)	0.976
Pearson R-square	34.84%
Likelihood Ratio Index	27.75%

6.4.1.3. Comparison of All Crashes "Before" and "After" Models

The final "before" and "after" models are negative binomial models since the "before" and "after" data display extra variations relative to Poisson distributions. All seven predictor variables are statistically significant in the "before" negative binomial model at a 20% significance level, while posted speed are not included in the "after" negative binomial model due to its insignificance. Regarding the goodness-of-fit, the "after" model performs better in explaining the variation of data than the "before" model based on either Pearson's R-square or likelihood ratio index.

The relative effects of each predictor variable in the final "before" and "after" models are in the same direction. The relative effect of ADT is 63% after signalization, higher than 32% before signalization, which means that intersection crashes are more sensitive to ADT changes after signalization. Intersections in urban areas would have about 10% more crashes than in rural areas, and intersections in business areas would have 20% more crashes than in other areas, regardless of before or after signalization. These results can be explained by the fact that the percentage of turning movements is relatively higher in urban areas and/or business areas. The average number of crashes would increase by 30% with the number of lanes increasing from less than or equal to four to more than four, regardless of before or after signalization. Generally more lanes on the roadway mean higher volume, which will cause more conflicts. For the variable posted speed in the case of before signalization, intersections with higher than 45 mph posted speed would have 16% less crashes as compared to intersections with lower or equal to 45 mph posted speed. This result may be explained by the speed differential that could be caused by the lower posted speed. From another perspective, drivers on the minor road may be more conservative and more careful to enter or cross the major road if the vehicles' speeds on this road are high. Nevertheless, traffic on the minor road will be assigned right-of-away after the installation of the traffic signal, and the effect of posted speed becomes insignificant. For the case of before signalization, intersections with median would have 30% more crashes than intersections without median, while this figure decreases to 20% for after signalization. This could be caused by the fact that roadways with median usually are serving high volume of traffic, at the same time, are more likely located in urban areas and business areas. Intersections with paved shoulder would have 15% less crashes as compared to unpaved shoulder for the case of before signalization and 10% less crashes as compared to unpaved shoulder for the case of after signalization. It is reasonable because paved shoulders generally can make drivers feel safer, provide space for right-turn vehicles, and could function as a travel lane in case vehicles run off road.

6.4.2. Models for Angle Crashes

6.4.2.1. The Model for Before Signalization

The "before" angle crash data had extra variation based on the results of initial Poisson model, with mean deviance equal to 1.634 and Pearson Chi-square ratio equal to 1.839. Then, Negative binomial regression was run with all seven predictor variables in the model. Of the seven-predictor variables, coefficients of location type, number of lanes and posted speed were insignificant at 20 percent significant level. Backward elimination method was used to remove the insignificant variables from the model. The estimated parameters of the negative binomial model for angle crashes for "before" are presented in Table 6.21. Goodness-of-fit statistics were also calculated for the negative binomial model and are shown in Table 6.22. The Pearson R-square and likelihood ratio index are 11.85% and 7.00% respectively indicating that the "before" model for angle crashes explain a relatively low percent of systematic variation of the data.

Table 6.21. Estimated Parameters of the Negative Binomial Model (Angle, Before)

Predictor Variable	Name	Level	Value	DOF	Coefficent Estimate	Standard Error	Relative Effect	Chi-square	Pr>Chisq
Intercept				1	-0.5629	0.1369		16.91	<0.0001
Average ADT	AVGADT	<15,000 vpd	0	1	0.2237	0.0859	1.25	6.79	0.0092
		15,000~30,000	1						
		>=30,000 vpd	2						
Surrounding Land Use	URBRUR	Urban	1	1	0.2829	0.1203	1.33	5.53	0.0187
		Rural	0						
Location Type	LOCATYPE	Business	1	~	~	~	~	~	~
		Other	0						
Number of Lanes on Major Road	LANE	> 4	1	~	~	~	~	~	~
		<= 4	0						
Posted Speed on Major Road	SPEED	> 45	1	~	~	~	~	~	~
		<=45	0						
Presence of Median on Major Road	MEDIAN	Divided	1	1	0.4183	0.1393	1.52	9.01	0.0027
		Undivided	0						
Shoulder Treatment	SHOULD	Paved	1	1	-0.3985	0.1377	0.67	8.38	0.0038
		Other	0						
Dispersion Parameter				1	0.4415	0.1036			

Table 6.22. Criteria for Assessing the Goodness-of-fit (Angle)

Item	Before (Poisson)	After (Negative Binomial)
Number of Observations (n)	436	437
Number of Predictor Variables in Model	4	4
Number of Parameters in Model (p)	4	4
Degree of Freedom (n-p-1)	431	432
Log-likelihood at Convergence	-558.35	-490.37
Restricted Log-likelihood	-600.36	-515.78
Deviance	422.55	404.67
Deviance/(n-p-1)	0.980	0.937
Pearson Chi-square	432.79	397.78
Pearson Chi-square/(n-p-1)	1.004	0.921
Pearson R-square	11.85%	16.63%
Likelihood Ratio Index	7.00%	4.93%

6.4.2.2. The Model for After Signalization

The calculated mean deviance and Pearson Chi-square ratio for the after model were 1.13 and 1.23 respectively, indicating that the Poisson regression is appropriate. Backward elimination method was used to take out those variables that were not significant at 20% significance level, including variables describing urban/rural, posted speed and shoulder. The other four variables were kept in the Poisson model. The estimated parameters of the Poisson model for "after" angle crashes are presented in Table 6.23. Goodness-of-fit statistics for the Poisson model were calculated and shown in Table 6.22.

6.4.2.3. Comparison of Angle Crashes "Before" and "After" Models

The “before” model is a negative binomial model because “before" angle crash data showed extra variation relative to Poisson distribution while the “after” model is a Poisson model because "after" angle crash data followed the Poisson distribution very well. In both models, the variable posted speed on major road is insignificant, which indicates that the change of posted speed from less than or equal to 45mph to higher has very limited effect on angle crash occurrence at intersections. Surrounding land use and shoulder treatment are significant in the "before" model but insignificant in the "after" model. Location type and number of lanes are insignificant in the "before" model but significant in the "after model. Presence of median is significant in both models.

Table 6.23. Estimated Parameters of the Poisson Model (Angle, After)

Predictor Variable	Name	Level	Value	DOF	Coefficent Estimate	Standard Error	Relative Effect	Chi-square	Pr>Chisq
Intercept				1	-0.8677	0.1639		28.03	<0.0001
Average ADT	AVGADT	<15,000 vpd	0	1	0.2924	0.0842	1.34	12.05	0.0005
		15,000~30,000	1						
		>=30,000 vpd	2						
Surrounding Land Use	URBRUR	Urban	1	~	~	~	~	~	~
		Rural	0						
Location Type	LOCATYPE	Business	1	1	0.1768	0.1391	1.19	1.62	0.2036
		Other	0						
Number of Lanes on Major Road	LANE	> 4	1	1	0.3654	0.1225	1.44	8.89	0.0029
		<= 4	0						
Posted Speed on Major Road	SPEED	> 45	1	~	~	~	~	~	~
		<=45	0						
Presence of Median on Major Road	MEDIAN	Divided	1	1	0.1698	0.1281	1.19	1.56	0.212
		Undivided	0						
Shoulder Treatment	SHOULD	Paved	1	~	~	~	~	~	~
		Other	0						

The Pearson R-square statistics and the likelihood ratio for both angle crash models are lower than for the models for all crashes. This happens very often to disaggregate data in modeling compared to aggregate data. It was also found that Pearson R-square statistic and likelihood ratio index are inconsistent in assessing the goodness-of-fit of developed models. The Pearson R-square increases from 11.85 % for the "before" model to 16.63% for the "after" model, while the likelihood ratio index decreases from 7.00% for the “before” model to 4.93% for the “after” model. Considering this finding, it is better to count on more than one statistic when evaluating the goodness-of-fit of generalized linear models.

6.4.3. Models for Left-turn Crashes

6.4.3.1. The Model for Before Signalization

Initial Poisson regression had mean deviance of 1.612 and Pearson Chi-square ratio of 1.939, indicating over-dispersion in the data. Therefore, negative binomial regression was performed. Of seven predictor variables, urban/rural and posted speed are insignificant at 50% significance level. Backward elimination method was used to remove these two variables and it was found that the model was insensitive to these two variables. The estimated parameters for the negative binomial model are presented in Table 6.24. The relative effects of each predictor variable in the negative binomial model for the “before” left-turn crashes are in the same direction as for all crashes. For example, with the increase of ADT from low to medium, there would be 26% more left-turn crashes. The estimated dispersion parameter is equal to 0.2717 indicating that crash data were only slightly over-dispersed. Table 6.25 shows the goodness-of-fit statistics calculated for the final negative binomial model. Pearson R-square is 13.84% and likelihood ratio index is 5.01% indicating that the model explains the variation very limitedly.

Table 6.24. Estimated Parameters of the Negative Binomial Model (Left-turn, Before)

Predictor Variable	Name	Level	Value	DOF	Coefficent Estimate	Standard Error	Relative Effect	Chi-Square	Pr>Chisq
Intercept				1	-0.7791	0.1694		21.16	<0.0001
Average ADT	AVGADT	<15,000 vpd	0	1	0.2297	0.0894	1.26	6.61	0.0102
		15,000~30,000	1						
		>=30,000 vpd	2						
Surrounding Land Use	URBRUR	Urban	1	~	~	~	~	~	~
		Rural	0						
Location Type	LOCATYPE	Business	1	1	0.3215	0.1518	1.38	4.49	0.0342
		Other	0						
Number of Lanes on Major Road	LANE	> 4	1	1	0.3304	0.1606	1.39	4.23	0.0397
		<= 4	0						
Posted Speed on Major Road	SPEED	> 45	1	~	~	~	~	~	~
		<=45	0						
Presence of Median on Major Road	MEDIAN	Divided	1	1	0.2156	0.1368	1.24	2.48	0.1149
		Undivided	0						
Shoulder Treatment	SHOULD	Paved	1	1	-0.2495	0.1330	0.78	3.52	0.0606
		Other	0						
Dispersion Parameter				1	0.2717	0.0970			

Table 6.25. Criteria for Assessing the Goodness-of-fit (Left-turn)

Item	Before (Poisson)	After (Negative Binomial)
Number of Observations (n)	434	436
Number of Predictor Variables in Model	5	5
Number of Parameters in Model (p)	5	5
Degree of Freedom (n-p-1)	428	430
Log-likelihood at Convergence	-516.833	-458.237
Restricted Log-likelihood	-544.098	-480.849
Deviance	423.225	392.697
Deviance/(n-p-1)	0.989	0.913
Pearson Chi-square	421.926	402.987
Pearson Chi-square/(n-p-1)	0.986	0.937
Pearson R-square	13.84%	15.90%
Likelihood Ratio Index	5.01%	4.70%

6.4.3.2. The Model for After Signalization

The Poisson model was run, and the "after" data showed only slightly over-dispersion with mean deviance equal to 1.205 and Pearson Chi-square ratio equal to 1.466. The variables urban/rural and posted speed were insignificant at 50% significance level. Then, backward elimination method was used to remove ported speed and urban/rural. The significance level of shoulder treatment increases form 22% to 17.35% once urban/rural and posted speed were removed from the model. The estimated parameters for the Poisson model are presented in Table 6.26. The relative effect of each predictor variable is in the same direction as in the “before” model. Goodness-of-fit statistics for the final model were calculated and presented in Table 6.25.

6.4.3.3. Comparison of Left-turn Crashes "Before" and "After" models

The "before" model is a negative binomial model because the "before" left-turn crash data were found to have extra variation relative to Poisson distribution. The "after" model is a Poisson model because the "after" left-turn crash data were found to match the Poisson distribution. For both models, posted speed and surrounding land use (urban/rural) were insignificant at 20 percent significance level. Other variables included in the after model affect the left-turn crashes in the same direction as in the “before” model. For example, keeping other characteristics constant, an intersection in business area would have more crashes than in other area whether before or after signalization. An

Table 6.26. Estimated Parameters of the Poisson Model (Left-turn, After)

Predictor Variable	Name	Level	Value	DOF	Coefficent Estimate	Standard Error	Relative Effect	Chi-square	Pr>Chisq
Intercept				1	-1.0985	0.1837		35.74	<0.0001
Average ADT	AVGADT	<15,000 vpd	0	1	0.1791	0.0888	1.20	4.07	0.0437
		15,000~30,000	1						
		>=30,000 vpd	2						
Surrounding Land Use	URBRUR	Urban	1	~	~	~	~	~	~
		Rural	0						
Location Type	LOCATYPE	Business	1	1	0.2286	0.1515	1.26	2.28	0.1312
		Other	0						
Number of Lanes on Major Road	LANE	> 4	1	1	0.2812	0.1343	1.32	4.38	0.0363
		<= 4	0						
Posted Speed on Major Road	SPEED	> 45	1	~	~	~	~	~	~
		<=45	0						
Presence of Median on Major Road	MEDIAN	Divided	1	1	0.5164	0.1452	1.68	12.65	0.0004
		Undivided	0						
Shoulder Treatment	SHOULD	Paved	1	1	-0.1675	0.1231	0.85	1.85	0.1735
		Other	0						

intersection with more than 4 lanes on the major road would have more crashes than one with less than or equal to 4 lanes whether before or after signalization. Regarding the goodness-of-fit of the two models, the Pearson R-square for the "after" model has a minor increase while the likelihood ratio index has a minor decrease. Overall, a similar conclusion can be made like for angle models that the regression models performed better for aggregate data that for disaggregate data.

6.4.4. Models for Rear-end Crashes

6.4.4.1. The Model for Before Signalization

The calculated mean deviance for the "before" Poisson regression with all predictor variables was 1.08 and Pearson Chi-square ratio was 1.13. In the model, variables describing surrounding land use and shoulder treatment were insignificant at 20 percent significance level. Backward elimination method was used to remove insignificant variables. Table 6.27 shows the estimated parameters for the "before" Poisson model. Finally, five variables were significant at 20 percent significance level. The relative effects of all variables except for presence of median have the same sign as in other developed models. Goodness-of-fit statistics for the "before" model were calculated and shown in Table 6.28. Pearson R-square of 30.54%, and likelihood ratio index of 10.16% show that the final Poisson model fit the data satisfactorily.

6.4.4.2. The Model for After Signalization

The after data show extra-variation relative to the Poisson model, with mean deviance equal to 1.875 and Pearson Chi-square ratio equal to 2.218. Therefore, negative binomial regression was used to overcome the over-dispersion. The variables describing surrounding land use, land use and posted speed were insignificant at 20 percent significance level. Backward elimination method was used to try different combinations of predictor variables and to remove the insignificant variables. Finally, only surrounding land use and location type were removed from the model, and the other five variables were included in the negative binomial model for rear-end crashes after signalization. Table 6.29 shows the estimated parameters. Goodness-of-fit statistics were also calculated and shown in Table 6.28.

Table 6.27. Estimated Parameters for the Poisson Model (Rear-end, Before)

Predictor Variable	Name	Level	Value	DOF	Coefficent Estimate	Standard Error	Relative Effect	Chi-square	Pr>Chisq
Intercept				1	-0.8804	0.1667		27.90	<0.0001
Average ADT	AVGADT	<15,000 vpd	0	1	0.5565	0.0811	1.74	47.08	<0.0001
		15,000~30,000	1						
		>=30,000 vpd	2						
Surrounding Land Use	URBRUR	Urban	1	~	~	~	~	~	~
		Rural	0						
Location Type	LOCATYPE	Business	1	1	0.2467	0.1482	1.28	2.77	0.0959
		Other	0						
Number of Lanes on Major Road	LANE	> 4	1	1	0.4203	0.1266	1.52	11.02	0.0009
		<= 4	0						
Posted Speed on Major Road	SPEED	> 45	1	1	-0.4708	0.1505	0.62	9.78	0.0018
		<=45	0						
Presence of Median on Major Road	MEDIAN	Divided	1	1	-0.1563	0.1204	0.86	1.68	0.1943
		Undivided	0						
Shoulder Treatment	SHOULD	Paved	1	~	~	~	~	~	~
		Other	0						

Table 6.28. Criteria for Assessing the Goodness-of-fit (Rear-end)

Item	Before (Poisson)	After (Negative Binomial)
Number of Observations (n)	441	440
Number of Predictor Variables in Model	6	5
Number of Parameters in Model (p)	6	5
Degree of Freedom (n-p-1)	434	434
Log-likelihood at Convergence	-504.805	-676.271
Restricted Log-likelihood	-561.864	-826.682
Deviance	468.978	429.635
Deviance/(n-p-1)	1.081	0.990
Pearson Chi-square	489.616	416.549
Pearson Chi-square/(n-p-1)	1.128	0.960
Pearson R-square	30.54%	35.74%
Likelihood Ratio Index	10.16%	18.19%

6.4.4.3. Comparison of Rear-end Crashes "Before" and "After" Models

Unlike the angle and left-turn crash data, the "before" rear-end crash data do not show extra-variation relative to Poisson model while the "after" data show extra-variation. The variable describing surrounding land use is insignificant at 20 percent significance level in both models. Shoulder treatment is insignificant in the "before" model but significant in the "after" model. Location type is significant in the "before" model but insignificant in the "after" model. The other variables are significant in both models. The relative effects of each predictor variable included in the models are in the same direction as in developed models for all crashes, except for the variable presence of median. The effect of median on rear-end crashes indicate that intersections would have less rear-end crashes if the major road is divided before signalization. This could be explained in the sense that the median could provide space for left-turn vehicles to clear the left through lane, otherwise, rear-end crashes are very likely to happen on the left through lane blocked by suddenly stopped left-turn vehicles. But, for signalized intersections, the presence of median would increase rear-end crashes which is a similar effect as in other models. Regarding the goodness-of-fit, both models have very high values of Pearson R-square and likelihood ratio index indicating that the developed models perform very well in explaining the systematic variations in the crash data. Relatively, the negative binomial model for the "after" data performs better than the Poisson model for the "before" data.

Table 6.29. Estimated Parameters of the Negative Binomial Model (Rear-end, After)

Predictor Variable	Name	Level	Value	DOF	Coefficent Estimate	Standard Error	Relative Effect	Chi-square	Pr>Chisq
Intercept				1	-0.2654	0.1095		5.87	0.0154
Average ADT	AVGADT	<15,000 vpd	0	1	0.6446	0.0724	1.91	79.30	<0.0001
		15,000~30,000	1						
		>=30,000 vpd	2						
Surrounding Land Use	URBRUR	Urban	1	~	~	~	~	~	~
		Rural	0						
Location Type	LOCATYPE	Business	1	~	~	~	~	~	~
		Other	0						
Number of Lanes on Major Road	LANE	> 4	1	1	0.1970	0.1074	1.22	3.36	0.0668
		<= 4	0						
Posted Speed on Major Road	SPEED	> 45	1	1	-0.1459	0.1081	0.86	1.82	0.1772
		<=45	0						
Presence of Median on Major Road	MEDIAN	Divided	1	1	0.2227	0.1088	1.25	4.19	0.0407
		Undivided	0						
Shoulder Treatment	SHOULD	Paved	1	1	-0.3267	0.0990	0.72	10.89	0.001
		Other	0						
Dispersion Parameter					0.2723	0.0549			

6.4.5. Models for Other Crashes

6.4.5.1. The Model for Before Signalization

The "before" other crash data fitted the Poisson model with mean deviance and Pearson Chi-square ratio equal to 0.91. Backward elimination method was used to remove insignificant variables which included location type, number of lanes on major road, and shoulder treatment. Finally, four variables that were significant at 20 percent significance level were included in the Poisson model. The results of the "before" model are shown in Table 6.30. The relative effects of predictor variables have the same sign as most of other developed models. For example, intersections with 50 mph posted speed on major road would have 35% less other type of crashes than intersections with 40 mph posted speed on major road. Intersections in urban areas would have 40% more other type of crashes than in rural areas. Also, goodness-of-fit statistics for the "before" model were calculated and shown in Table 6.31. Pearson R-square and likelihood ratio index indicate that the developed model has limited ability in explaining the data variation.

6.4.5.2. The Model for After Signalization

The data show no extra-variation relative to Poisson model. Backward elimination method was used to remove the insignificant predictor variables that include land use, posted speed, median, and shoulder. The results of the Poisson model are shown in Table 6.32. Only ADT, urban/rural and number of lanes were included in the "after" model. Goodness-of-fit statistics of the model were also calculated and are shown in Table 6.31.

6.4.5.3. Comparison of Other Crashes "Before" and "After" Models

It was found that both the "before" and "after" data for other crashes follow the Poisson distribution without showing extra variation. Thus, Poisson regression is an appropriate choice in both cases. Variables describing location type and shoulder treatment are insignificant in both Poisson models. Number of lanes is insignificant in the "before" model, while posted speed and median are insignificant in the "after" model. The variables included in the models have relative effects in the same direction of other models. Goodness-of-fit statistics show that the "after" model fits the "after" data better than the "before" model fits the "before" data.

Table 6.30. Estimated Parameters for the Poisson Model (All Other, Before)

Predictor Variable	Name	Level	Value	DOF	Coefficent Estimate	Standard Error	Relative Effect	Chi-square	Pr>Chisq
Intercept				1	-0.8027	0.1287		38.87	<0.0001
Average ADT	AVGADT	<15,000 vpd	0	1	0.3350	0.0748	1.40	20.08	<0.0001
		15,000~30,000	1						
		>=30,000 vpd	2						
Surrounding Land Use	URBRUR	Urban	1	1	0.3408	0.1051	1.41	10.52	0.0012
		Rural	0						
Location Type	LOCATYPE	Business	1	~	~	~	~	~	~
		Other	0						
Number of Lanes on Major Road	LANE	> 4	1	~	~	~	~	~	~
		<= 4	0						
Posted Speed on Major Road	SPEED	> 45	1	1	-0.4362	0.1422	0.65	9.41	0.0022
		<=45	0						
Presence of Median on Major Road	MEDIAN	Divided	1	1	0.2727	0.1422	1.31	4.81	0.0284
		Undivided	0						
Shoulder Treatment	SHOULD	Paved	1	~	~	~	~	~	~
		Other	0						

Table 6.31. Criteria for Assessing the Goodness-of-fit (All Other)

Item	Before (Poisson)	After (Negative Binomial)
Number of Observations (n)	440	437
Number of Predictor Variables in Model	4	3
Number of Parameters in Model (p)	4	3
Degree of Freedom (n-p-1)	435	433
Log-likelihood at Convergence	-491.213	-495.212
Restricted Log-likelihood	-521.98	-540.979
Deviance	397.06	336.046
Deviance/(n-p-1)	0.913	0.776
Pearson Chi-square	399.097	317.303
Pearson Chi-square/(n-p-1)	0.917	0.733
Pearson R-square	17.61%	31.43%
Likelihood Ratio Index	5.89%	8.46%

6.5. Impacts of Signalization on Crashes

Once the models for each crash type considered were developed, the average numbers of crashes before and after signalization were estimated for all crashes and for each crash type for intersections with different characteristics. These characteristics varied accordingly with the variables used: ADT, surrounding land use, location type, number of lanes, posted speed, median, and shoulder type. The impacts of signalization on intersection crashes were estimated by subtracting the average number of crashes after signalization from the number of crashes before signalization. A positive sign indicates an increase in crashes and a negative sign a decrease. These impacts were estimated with the average number of crashes obtained from the developed models and not from crash history in order to make a comparison between estimated values for both the before and after period, and not from a combination of values from real data for the before period and predicted results for the after period. Tables were calculated for the impacts of signalization on crashes for all crashes and each crash type. Appendix C presents the results for all crashes, angle crashes, left-turn crashes, rear-end crashes and all other crashes for different characteristics of intersections.

Table 6.32. Estimated Parameters of the Poisson Model (All Other, After)

Predictor Variable	Name	Level	Value	DOF	Coefficent Estimate	Standard Error	Relative Effect	Chi-square	Pr>Chisq
Intercept				1	-0.8928	0.1208		54.64	<0.0001
Average ADT	AVGADT	<15,000 vpd	0	1	0.5191	0.0745	1.68	48.58	<0.0001
		15,000~30,000	1						
		>=30,000 vpd	2						
Surrounding Land Use	URBRUR	Urban	1	1	0.3006	0.0974	1.35	9.53	0.002
		Rural	0						
Location Type	LOCATYPE	Business	1	~	~	~	~	~	~
		Other	0						
Number of Lanes on Major Road	LANE	> 4	1	1	0.2390	0.1118	1.27	4.57	0.0325
		<= 4	0						
Posted Speed on Major Road	SPEED	> 45	1	~	~	~	~	~	~
		<=45	0						
Presence of Median on Major Road	MEDIAN	Divided	1	~	~	~	~	~	~
		Undivided	0						
Shoulder Treatment	SHOULD	Paved	1	~	~	~	~	~	~
		Other	0						

These estimated impacts of signalization on crashes for intersections with different characteristics would provide information in regard to the increase or decrease in the number of all crashes and different type of crashes, and might help traffic engineers when considering the installation of a traffic signal as a solution for the intersection problems. The estimated impacts are not to be used as an instrument to decide but rather as a tool to evaluate how a signal may affect crashes at an intersection under analysis. In order to estimate the impacts of signalization on intersection crashes, the engineer could either use the tables included in the appendices or the software provided with this report. For the tables, each table has specific characteristics of the intersections for which those results were estimated, and for the software, the specific characteristics could be used as input in order to estimate the impacts of signalization.

6.6. Model Validation

A group of 30 newly signalized intersections were not considered in the intersection modeling database in order to compare the results obtained from the developed models to the real data values. Based on the characteristics of the intersections and the models for all crashes, the number of crashes was determined for the before and after period for each one of the intersections. Then, an average of the estimated number of all crashes was done and compare to the average of the number of all crashes for the actual data for both periods. The difference between the averages for either the before or after signalization periods was very small. The following table shows these results.

Table 6.33. Average Number of All Crashes
for Actual Data and Predicted Values

Period	Average Number of All Crashes	
	Actual	Predicted
Before	3.55	3.80
After	4.53	4.50

CHAPTER 7: RESULTS FOR OPERATIONAL RESEARCH MODELING

7.1. Crash Prediction

CBCP was used to predict total number of all crashes, angle crashes, rear-end crashes, left-turn crashes, and “all other” crashes, which includes the remainder of crashes, at intersections before and after traffic signal installation. For this study, the impact of traffic signal installation on intersection crashes was computed using the predicted crash frequencies before and after signalization. Even though, the CBCP procedure could directly predict this impact, it was not used because the method is based on known data, and its prediction accuracy is closely dependent on the availability of data. In the research, with all the seven characteristic variables that define an individual intersection, few data points for each unique intersection type would be available. Therefore, the statistical regression analysis, specifically, the lognormal modeling using the predicted crashes was applied to estimate the impacts of signal installation on crashes at intersections.

7.2. Lognormal Modeling

Lognormal modeling was applied using the predictive number of crashes for all crashes and each crash type estimated with CBCP. The following paragraphs present the results of this modeling.

7.2.1. Models for All Crashes

7.2.1.1. The Model for Before Signalization

In particular, residual analysis was used to check the model fit. In general, the best-fitted data points should have residuals whose absolute values are close to zero. The lognormal regression results are shown Tables 7.1 and 7.2, and in Figure 7.1. As it is expected, all residuals are between 1 and –1 and most are close to 0. This implies that the model fits the data satisfactorily. The p-value in Table 7.1 means that the regression model is significant at more than 99% confidence level.

Table 7.1. Estimated Parameters of Lognormal Model (All, Before)

Predictor Variable	Name	Level	Value	DOF	Coefficient Estimate	Standard Error	Relative Effect	t-Stat	P-value
Intercept				1	0.875	0.045		19.320	<=0.001
Average ADT	AVGADT	< 15,000 vpd	0	2	0.120	0.029	1.127	4.097	<=0.001
		15,000~30,000	1						
		>= 30,000 vpd	2						
Urban/Rural	URBRUR	Urban	1	1	0.190	0.032	1.210	6.018	<=0.001
		Rural	0						
Land Use of Surrounding Area	LOCATYPE	Business	1	1	0.038	0.038	1.038	0.998	0.319
		Other	0						
Number of Lanes on Major Road	LANE	> 4	1	1	0.310	0.050	1.363	6.211	<=0.001
		<= 4	0						
Posted Speed on Major Road	SPEED	> 45	1	1	-0.184	0.039	0.832	-4.757	<=0.001
		<= 45	0						
Presence of Median on Major Road	MEDIAN	Divided	1	1	0.400	0.034	1.491	11.751	<=0.001
		Undivided	0						
Shoulder Treatment	SHOULD	Paved	1	1	-0.099	0.035	0.905	-2.828	0.005
		Other	0						

Table 7.2. Analysis of Variance (Total, Before)

Square of Variation	Sum of Squares	DOF	Mean Square	F-Stat	P-Value
Regression	44.101	7	6.30	63.14	< 0.001
Error	44.001	441	0.10		
Total	88.103	448			

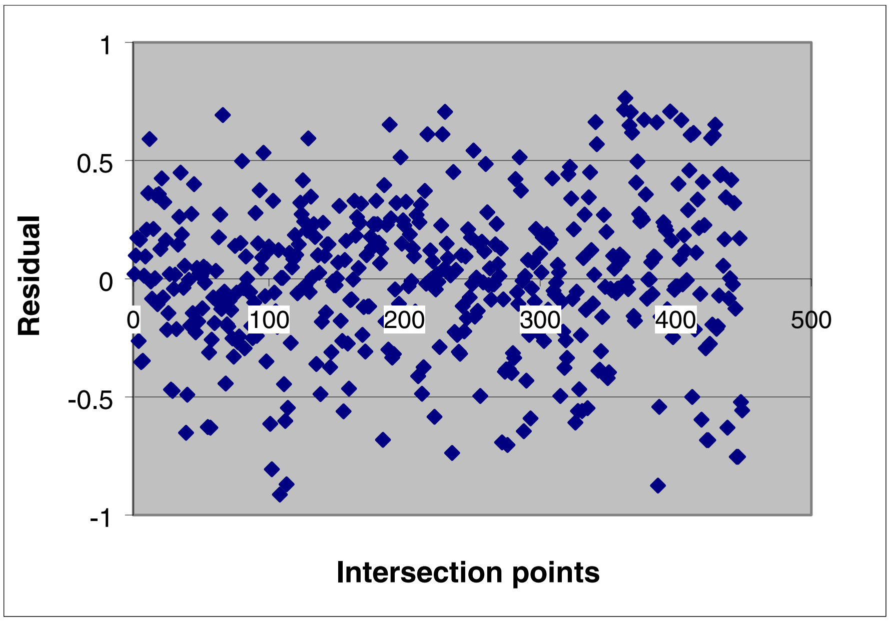

Figure 7.1. Residual for Lognormal Model (Total, Before)

From Table 7.1, it can be seen clearly that all the estimated parameters (except LOCALTYPE) were significant at more than 99 % confidence level, while the estimated parameter for LOCALTYPE is significant at only 68 % confident level. The variable LOCALTYPE is included in the "before" model to be consistent with the statistical model presented in Chapter 6. The contents in Table 7.1 are the same as those in Table 6.21, except the last two columns, which are explained in Table 7.3.

The relative effect of ADT is 1.127. This means that the average number of crashes would increase by 12.7% if the ADT increases from the low level to the medium level given that all the other variables are unchanged. Similarly, if ADT increases from the medium level to the high level, the average number of crashes would increase by 12.7%.

Intersections in urban areas would have 21 % more crashes than similar intersections in rural areas. Intersections in business area would have about 4% more crashes than similar ones in other areas. Intersections with median on the major road would have 49% more crashes than similar ones without median on the major road. Similarly, more lanes means more crashes. However, high-speed intersections (mph > 45) have about 17% fewer crashes than low speed intersections (mph ≤ 45). Likewise, intersections with paved shoulder are safer than those with other types of shoulder. All these results are consistent with the results of the statistical models.

Table 7.3. Explanation of Contents

Column	*Explanation*
t-Stat	t-test statistic for testing whether the parameter is 0. t is computed as the ratio of the point estimate of a parameter to the standard error of the parameter. Note that this is usually a partial or marginal test, because the point estimate of a parameter depends on all the other regressor variables that are in the regression model.
P-value	The probability of obtaining a t-test statistic greater than the obtained, given that the true parameter is 0. A small P-value indicates that the true parameter is not 0.

7.2.1.2. The Model for After Signalization

The "after" model for all crashes was developed through a similar procedure as the "before" model. In particular, all the seven variables were initially included and a lognormal regression was performed. It turned out that variable SPEED is insignificant. With variable SPEED removed, the lognormal regression was run again. As the "before" model, the statistics indicate that the lognormal regression is an appropriate choice. The obtained parameter estimates along with statistics and variances are presented in Tables 7.4 and 7.5, respectively. Residuals were also computed and plotted in Figure 7.2. Again, all residuals are between 1 and –1, and most are close to 0. This indicates the model fits the data satisfactorily.

The P-values in Table 7.4 indicates that all the estimated parameters included in the model were significant at more than 99 % confidence level.

Table 7.4. Estimated Parameters of Final Lognormal Model (All, After)

Predictor Variable	Name	Level	Value	DOF	Coefficient Estimate	Standard Error	Relative Effect	t-Stat	P-value
Intercept				1	0.679	0.042		16.135	<=0.001
Average ADT	AVGADT	< 15,000 vpd	0	2	0.260	0.026	1.297	10.090	<=0.001
		15,000~30,000	1						
		>= 30,000 vpd	2						
Urban/Rural	URBRUR	Urban	1	1	0.148	0.030	1.160	4.878	<=0.001
		Rural	0						
Land Use of Surrounding Area	LOCATYPE	Business	1	1	0.184	0.036	1.202	5.149	<=0.001
		Other	0						
Number of Lanes on Major Road	LANE	> 4	1	1	0.204	0.042	1.227	4.921	<=0.001
		<= 4	0						
Posted Speed on Major Road	SPEED	> 45	1	1	~	~	~	~	~
		<= 45	0						
Presence of Median on Major Road	MEDIAN	Divided	1	1	0.428	0.033	1.534	12.970	<=0.001
		Undivided	0						
Shoulder Treatment	SHOULD	Paved	1	1	-0.109	0.032	0.897	-3.374	0.001
		Other	0						

Table 7.5. Analysis of Variance (Total, After)

Square of Variation	Sum of Squares	DOF	Mean Square	F-Stat	P-Value
Regression	66.250	6	11.042	117.303	< 0.001
Error	41.699	442	0.094		
Total	107.979	448			

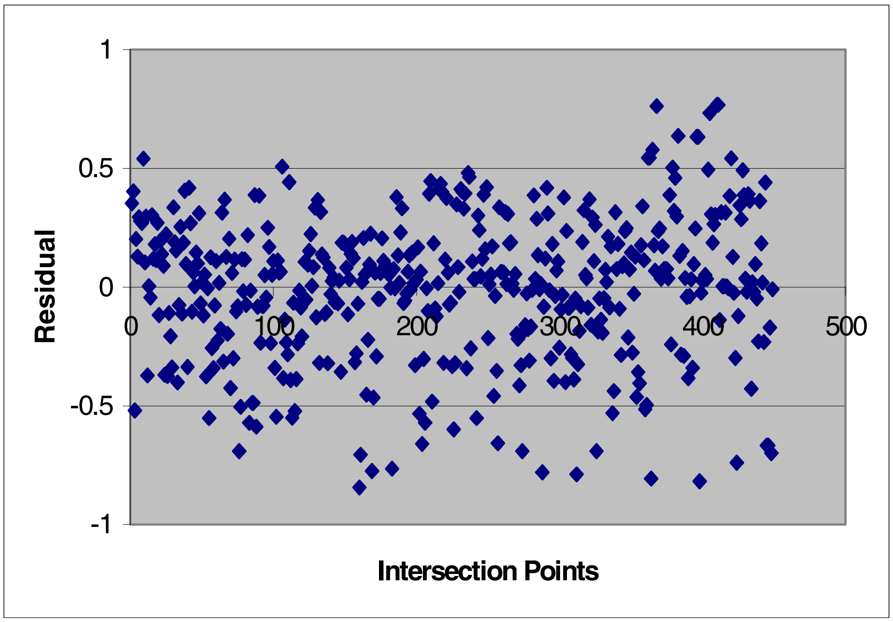

Figure 7.2. Residual for Lognormal Model (Total, After)

The relative effect of ADT is 1.297. This means that the average number of crashes would increase by 29.7% if the ADT increases by one level (i.e., from the low level to the medium level, or from the medium level to the high level), given that all the other variables are unchanged. Urban area intersections would have 16% more crashes than similar rural area intersections. Intersections in business area would have about 20% more crashes than similar ones in other areas. Intersections with median on the major road would have about 53% more crashes than similar ones without median on the major road. Similarly, more lanes means more crashes. As mentioned earlier, speed is insignificant and is excluded from the model. Intersections with paved shoulder are about 9% safer than those with other types of shoulder, in terms of number of all crashes.

7.2.1.3. Comparison of All Crashes "Before" and "After" Models

The relative effects of each model variable, in the "before" and "after" models, are in the same direction. The relative effect of ADT is 1.127 and 1.297 for the "before" and "after" models, respectively. This implies that intersection crashes are more sensitive to ADT changes after signalization. Similar conclusion can be observed for LOCALTYPE. In other words, variable LOCALTYPE is more significant after signalization. The variables LANE and SPEED are just the opposite. That is, intersection crashes are more sensitive to LANE or SPEED changes before signalization. In fact, intersection crashes are insensitive to SPEED after signalization. The relative effects of URBRUR on "before" and "after" crashes are pretty the same, indicating that intersection crashes are not very sensitive to whether intersections are in an urban area or in a rural area. Similar conclusions can be derived for MEDIAN and SHOULDER.

7.2. 2. Models for Angle Crashes

7.2.2.1. The Model for Before Signalization

Initial run of the lognormal regression included all the seven variables. Run results indicated that parameters for four variables AVGADT, SPEED, LANE and LOCALTYPE are insignificant. However, only three variables SPEED, LANE and LOCALTYP have been removed from the model. Variable AVGADT is included in the model in order to be consistent with the statistical model. The model with the remaining four variables had most residuals closed to zero, as in the case of total crashes, indicating that the model fits the data satisfactorily. The regression results are presented in Tables 7.6 and 7.7.

Table 7.6. Estimated Parameters of Lognormal Model (Angle, Before)

Predictor Variable	Name	Level	Value	DOF	Coefficient Estimate	Standard Error	Relative Effect	t-Stat	P-value
Intercept				1	-0.304	0.050		-6.117	<=0.001
Average ADT	AVGADT	< 15,000 vpd	0	2	0.023	0.040	1.023	0.574	0.566
		15,000~30,000	1						
		>= 30,000 vpd	2						
Urban/Rural	URBRUR	Urban	1	1	0.218	0.046	1.243	4.713	<=0.001
		Rural	0						
Land Use of Surrounding Area	LOCATYPE	Business	1	1	~	~	~	~	~
		Other	0						
Number of Lanes on Major Road	LANE	> 4	1	1	~	~	~	~	~
		<= 4	0						
Posted Speed on Major Road	SPEED	> 45	1	1	~	~	~	~	~
		<= 45	0						
Presence of Median on Major Road	MEDIAN	Divided	1	1	0.523	0.050	1.686	10.350	<=0.001
		Undivided	0						
Shoulder Treatment	SHOULD	Paved	1	1	-0.379	0.051	0.684	-7.453	<=0.001
		Other	0						

Table 7.7. Analysis of Variance (Angle, Before)

Square of Variation	Sum of Squares	DOF	Mean Square	F-Stat	P-Value
Regression	41.595	4	10.399	37.279	< 0.001
Error	96.775	440	0.22		
Total	138.37	444			

7.2.2.2. The Model for After Signalization

Initial run of the lognormal regression included all the seven variables. Run results indicated that parameters for three variables SPEED, SHOULD and URBRUR are very insignificant, and these three variables should be removed from the model. As in the previous cases, most residuals are closed to zero for the lognormal model with four variables, indicating that the model fits the data satisfactorily. The regression results are presented in Tables 7.8 and 7.9

7.2.2.3. Comparison of Angle Crashes "Before" and "After" Models

In both the "before" and "after" models, medium is the only variable that is extreme significant, and SPEED is the only variable that is insignificant, Variables URBRUR and SHOULD are both significant in the "before" model but insignificant in the "after" model. To the contrary, variables LOCALTYPE and LANE are both insignificant in the "before" model but significant in the "after" model. Variable ADT is significant at more than 99% confidence level in the "after" model, but at pretty low (43%) confidence level in the "before" model.

Table 7.8. Estimated Parameters of Lognormal Model (Angle, After)

Predictor Variable	Name	Level	Value	DOF	Coefficient Estimate	Standard Error	Relative Effect	t-Stat	P-value
Intercept				1	-0.757	0.051		-14.85	<=0.001
Average ADT	AVGADT	< 15,000 vpd	0	2	0.220	0.035	1.247	6.370	<=0.001
		15,000~30,000	1						
		>= 30,000 vpd	2						
Urban/Rural	URBRUR	Urban	1	1	~	~	~	~	~
		Rural	0						
Land Use of Surrounding Area	LOCATYPE	Business	1	1	0.234	0.044	1.264	5.264	<=0.001
		Other	0						
Number of Lanes on Major Road	LANE	> 4	1	1	0.238	0.051	1.269	4.649	<=0.001
		<= 4	0						
Posted Speed on Major Road	SPEED	> 45	1	1	~	~	~	~	~
		<= 45	0						
Presence of Median on Major Road	MEDIAN	Divided	1	1	0.201	0.041	1.222	4.956	<=0.001
		Undivided	0						
Shoulder Treatment	SHOULD	Paved	1	1	~	~	~	~	~
		Other	0						

Table 7.9. Analysis of Variance (Angle, After)

Square of Variation	Sum of Squares	DOF	Mean Square	F-Stat	P-Value
Regression	42.833	4	10.71	35.365	< 0.001
Error	132.622	438	0.30		
Total	175.455	442			

7.2.3. Models for Left-turn Crashes

7.2.3.1. The Model for Before Signalization

Initial run of the lognormal regression included all the seven variables. Run results indicated that two variables SPEED and URBRUR are very insignificant, and are removed from the model. The model with the remaining five variables was rerun, and again, most residuals are closed to zero, indicating the model fits the data satisfactorily. The regression results are presented in Tables 7.10 and 7.11. All the remaining variables are significant at more than 98% confidence level.

Table 7.10. Estimated Parameters of Lognormal Model (Left-turn, Before)

Predictor Variable	Name	Level	Value	DOF	Coefficient Estimate	Standard Error	Relative Effect	t-Stat	P-value
Intercept				1	-0.664	0.063		-10.50	<=0.001
Average ADT	AVGADT	< 15,000 vpd	0	2	0.178	0.047	1.195	3.825	<=0.001
		15,000~30,000	1						
		>= 30,000 vpd	2						
Urban/Rural	URBRUR	Urban	1	1	~	~	~	~	~
		Rural	0						
Land Use of Surrounding Area	LOCATYPE	Business	1	1	0.192	0.055	1.212	3.482	0.001
		Other	0						
Number of Lanes on Major Road	LANE	> 4	1	1	0.351	0.076	1.421	4.638	<=0.001
		<= 4	0						
Posted Speed on Major Road	SPEED	> 45	1	1	~	~	~	~	~
		<= 45	0						
Presence of Median on Major Road	MEDIAN	Divided	1	1	0.405	0.051	1.500	7.915	<=0.001
		Undivided	0						
Shoulder Treatment	SHOULD	Paved	1	1	-0.128	0.052	0.880	-2.474	0.014
		Other	0						

Table 7.11. Analysis of Variance (Left-turn, Before)

Square of Variation	Sum of Squares	DOF	Mean Square	F-Stat	P-Value
Regression	48.244	5	9.65	41.857	< 0.001
Error	100.967	438	0.231		
Total	149.211	443			

7.2.3.2. The Model for After Signalization

Initial run of the lognormal regression included all the seven variables. Run results indicated that the same two variables SPEED and URBRUR are very insignificant, and are removed from the model. The lognormal regression was run again. As in the previous models, most residuals are closed to zero, indicating the model fits the data satisfactorily. The regression results are presented in Tables 7.12 and 7.13. It can be seen from Table 7.13 that LOCAL TYPE is significant at about 80% confidence level, while the other 4 remaining parameters are significant at more than 99% confidence level.

7.2.3.3. Comparison of Left-turn Crashes "Before" and "After" Models

Variables URBRUR and SPEED are very insensitive both before and after signalization, and were excluded from the models. All the remaining model parameters are pretty significant. In terms of relative effect, median has the largest impact on intersection crashes both before and after signalization. In particular, intersections with median on the major road would have 50% and 92.4% more crashes than those without median before and after signalization, respectively. Left-turn intersection crashes are less sensitive to ADT, MEDIUM and SHOULD but more sensitive to LOCATYPE and LANE before signalization than after signalization. It is worth pointing out that intersections with paved shoulder would have 12% and 24% fewer left-turn crashes than similar ones with other type of shoulder, before and after signalization, respectively.

Table 7.12. Estimated Parameters of Lognormal Model (Left-turn, After)

Predictor Variable	Name	Level	Value	DOF	Coefficient Estimate	Standard Error	Relative Effect	t-Stat	P-value
Intercept				1	-1.101	0.080		-13.73	<=0.001
Average ADT	AVGADT	< 15,000 vpd	0	2	0.271	0.056	1.311	4.815	<=0.001
		15,000~30,000	1						
		>= 30,000 vpd	2						
Urban/Rural	URBRUR	Urban	1	1	~	~	~	~	~
		Rural	0						
Land Use of Surrounding Area	LOCATYPE	Business	1	1	0.088	0.070	1.092	1.263	0.207
		Other	0						
Number of Lanes on Major Road	LANE	> 4	1	1	0.313	0.081	1.367	3.861	<=0.001
		<= 4	0						
Posted Speed on Major Road	SPEED	> 45	1	1	~	~	~	~	~
		<= 45	0						
Presence of Median on Major Road	MEDIAN	Divided	1	1	0.654	0.063	1.924	10.439	<=0.001
		Undivided	0						
Shoulder Treatment	SHOULD	Paved	1	1	-0.274	0.061	0.760	-4.529	<=0.001
		Other	0						

Table 7.13. Analysis of Variance (Left-turn, After)

Square of Variation	Sum of Squares	DOF	Mean Square	F-Stat	P-Value
Regression	103.328	5	20.67	61.212	< 0.001
Error	145.17	430	0.34		
Total	248.498	435			

7.2.4. Models for Rear-end Crashes

7.2.4.1. The Model for Before Signalization

Initial run of the lognormal regression included all the seven variables. Run results indicated that two variables SHOULD and URBRUR are very insignificant, and are removed from the model. The regression was run again and results are presented in Tables 7.14 and 7.15. All the remaining variables are significant at more than 99% confidence level, except that the confidence level for MEDIAN is only 70%. This variable was kept in the model in order to be consistent with the statistical model for before signalization for rear-end crashes.

7.2.4.2. The Model for After Signalization

Initial run of the lognormal regression included all the seven variables. Run results indicated that two variables URBRUR and LOCATYPE are insignificant, and were removed from the model. The lognormal regression was run again with the remaining five variables. The regression results are presented in Tables 7.16 and 7.17. All the remaining variables are significant with a confidence level of at least 80 %. In fact, all the remaining model parameters are at a confidence level of at least 97%, except for SPEED with an 84% confidence level.

Table 7.14. Estimated Parameters of Lognormal Model (Rear-end, Before)

Predictor Variable	Name	Level	Value	DOF	Coefficient Estimate	Standard Error	Relative Effect	t-Stat	P-value
Intercept				1	-0.910	0.060		-15.07	<=0.001
Average ADT	AVGADT	< 15,000 vpd	0	2	0.507	0.040	1.660	12.716	<=0.001
		15,000~30,000	1						
		>= 30,000 vpd	2						
Urban/Rural	URBRUR	Urban	1	1	~	~	~	~	~
		Rural	0						
Land Use of Surrounding Area	LOCATYPE	Business	1	1	0.345	0.052	1.413	6.674	<=0.001
		Other	0						
Number of Lanes on Major Road	LANE	> 4	1	1	0.201	0.068	1.223	2.960	0.003
		<= 4	0						
Posted Speed on Major Road	SPEED	> 45	1	1	-0.338	0.052	0.713	-6.453	<=0.001
		<= 45	0						
Presence of Median on Major Road	MEDIAN	Divided	1	1	-0.049	0.047	0.952	-1.041	0.298
		Undivided	0						
Shoulder Treatment	SHOULD	Paved	1	1	~	~	~	~	~
		Other	0						

Table 7.15. Analysis of Variance (Rear-end, Before)

Square of Variation	Sum of Squares	DOF	Mean Square	F-Stat	P-Value
Regression	84.252	5	16.85	87.20	< 0.001
Error	85.605	443	0.19		
Total	169.857	448			

7.2.4.3. Comparison of Rear-end Crashes "Before" and "After" Models

Variable URBRUR is very insignificant in both the "before' and "after" models and is deleted from consideration. All the remaining variables are pretty significant. ADT has significant relative effects both before and after signalization. Specifically, number of rear-end intersection crashes would increase by almost 60% both before and after signalization as ADT increases from low to medium level or from medium to high level. One observation that is worth noting is that before signalization, intersections with median on the major road have about 5% fewer rear-end crashes than similar intersections without median on the major road, but after signalization, intersections with median on the major road have 65.7% more rear-end crashes than similar intersections without median on the major road. Rear-end intersection crashes are very sensitive to LOCATYPE before signalization, but insensitive to LOCATYPE after signalization.

Table 7.16. Estimated Parameters of Lognormal Model (Rear-end, After)

Predictor Variable	Name	Level	Value	DOF	Coefficient Estimate	Standard Error	Relative Effect	t-Stat	P-value
Intercept				1	-0.305	0.043		-7.092	<=0.001
Average ADT	AVGADT	< 15,000 vpd	0	2	0.461	0.034	1.586	13.504	<=0.001
		15,000~30,000	1						
		>= 30,000 vpd	2						
Urban/Rural	URBRUR	Urban	1	1	~	~	~	~	~
		Rural	0						
Land Use of Surrounding Area	LOCATYPE	Business	1	1	~	~	~	~	~
		Other	0						
Number of Lanes on Major Road	LANE	> 4	1	1	0.122	0.055	1.130	2.211	0.028
		<= 4	0						
Posted Speed on Major Road	SPEED	> 45	1	1	0.066	0.046	1.068	1.425	0.155
		<= 45	0						
Presence of Median on Major Road	MEDIAN	Divided	1	1	0.505	0.045	1.657	11.343	<=0.001
		Undivided	0						
Shoulder Treatment	SHOULD	Paved	1	1	-0.294	0.042	0.745	-7.082	<=0.001
		Other	0						

Table 7.17. Analysis of Variance (Rear-end, After)

Square of Variation	Sum of Squares	DOF	Mean Square	F-Stat	P-Value
Regression	115.241	5	23.048	138.965	< 0.001
Error	73.640	443	0.166		
Total	188.881	448			

7.2.5. Models for Other Crashes

7.2.5.1. The Model for Before Signalization

As usual, the initial run of the lognormal regression included all the seven variables. Run results indicated that three variables LOCALTYPE, LANE and SHOULD are very insignificant, and have been removed from the model. The lognormal regression was run again and results are presented in Tables 7.18 and 7.19. All the remaining parameters are significant at more than 93% confidence level.

7.2.5.2. The Model for After Signalization

Again, the initial run of the lognormal regression included all the model parameters. Run results indicated that four variables LOCALTYPE, SPEED, MEDIUM and SHOULD are insignificant, and are deleted from the model. The lognormal regression was run again with the remaining four variables. The regression results are presented in Tables 7.20 and 7.21. All the remaining model parameters are significant with a confidence level of at least 99 %.

Table 7.18. Estimated Parameters of Lognormal Model (Other, Before)

Predictor Variable	Name	Level	Value	DOF	Coefficient Estimate	Standard Error	Relative Effect	t-Stat	P-value
Intercept				1	-0.760	0.045		-16.99	<=0.001
Average ADT	AVGADT	< 15,000 vpd	0	2	0.377	0.034	1.457	11.098	<=0.001
		15,000~30,000	1						
		>= 30,000 vpd	2						
Urban/Rural	URBRUR	Urban	1	1	0.326	0.040	1.386	8.258	<=0.001
		Rural	0						
Land Use of Surrounding Area	LOCATYPE	Business	1	1	~	~	~	~	~
		Other	0						
Number of Lanes on Major Road	LANE	> 4	1	1	~	~	~	~	~
		<= 4	0						
Posted Speed on Major Road	SPEED	> 45	1	1	-0.146	0.047	0.865	-3.071	0.002
		<= 45	0						
Presence of Median on Major Road	MEDIAN	Divided	1	1	0.080	0.043	1.083	1.838	0.067
		Undivided	0						
Shoulder Treatment	SHOULD	Paved	1	1	~	~	~	~	~
		Other	0						

Table 7.19. Analysis of Variance (Other, Before)

Square of Variation	Sum of Squares	DOF	Mean Square	F-Stat	P-Value
Regression	45.858	4	11.464	70.121	< 0.001
Error	71.774	439	0.163		
Total	117.632	443			

7.2.5.3. Comparison of Other Crashes "Before" and "After" Models

Two variables LOCALTYPE and SHOULD are very insignificant in both the "before' and "after" models and are both deleted from consideration. Two variables ADT and URBRUR are significant in both the "before" and "after" models. Variables SPEED and MEDIUM are significant before signalization but insignificant after signalization. To the contrary, variable LANE is insignificant before signalization but significant after signalization. ADT and URBRUR have the largest relative effects on intersection crashes both before and after signalization.

Table 7.20. Estimated Parameters of Lognormal Model (Other, After)

Predictor Variable	Name	Level	Value	DOF	Coefficient Estimate	Standard Error	Relative Effect	t-Stat	P-value
Intercept				1	-0.697	0.042		-16.47	<=0.001
Average ADT	AVGADT	< 15,000 vpd	0	2	0.361	0.033	1.435	10.860	<=0.001
		15,000~30,000	1						
		>= 30,000 vpd	2						
Urban/Rural	URBRUR	Urban	1	1	0.346	0.038	1.414	9.196	<=0.001
		Rural	0						
Land Use of Surrounding Area	LOCATYPE	Business	1	1	~	~	~	~	~
		Other	0						
Number of Lanes on Major Road	LANE	> 4	1	1	0.179	0.054	1.196	3.305	0.001
		<= 4	0						
Posted Speed on Major Road	SPEED	> 45	1	1	~	~	~	~	~
		<= 45	0						
Presence of Median on Major Road	MEDIAN	Divided	1	1	~	~	~	~	~
		Undivided	0						
Shoulder Treatment	SHOULD	Paved	1	1	~	~	~	~	~
		Other	0						

Table 7.21. Analysis of Variance (Other, After)

Square of Variation	Sum of Squares	DOF	Mean Square	F-Stat	P-Value
Regression	50.982	3	16.994	110.326	< 0.001
Error	67.622	439	0.154		
Total	118.604	442			

7.3. Impacts of Signalization on Crashes

Similar to the impacts of signalization on crashes for statistical models, the average numbers of crashes before and after signalization were estimated for all crashes and for each crash type for intersections with different characteristics based on lognormal models. Tables were calculated for the impacts of signalization on crashes for all crashes and each crash type. Appendix D presents the results for all crashes, angle crashes, left-turn crashes, rear-end crashes, and all other crashes for different characteristics of intersections.

Furthermore, tables with the average impacts of signalization on intersection crashes were calculated and are presented in Appendix E. These average impacts were estimated by averaging the statistical modeling and operational research modeling impacts of signalization on intersection crashes presented on Appendices C and D.

CHAPTER 8. SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

8.1. Summary

This research was performed to evaluate the impacts of signalization on traffic crashes at intersections. A ten-year crash history database including traffic crashes from intersections all over Florida was used for the research. In the first phase of the research, a before-and-after analysis was performed to compare number of crashes and crash rates based on different types, severities and surrounding land uses. Distribution fitting for the Poisson distribution or the Negative Binomial distribution was performed based on crash data. The 50th and 85th percentile values were estimated from the distribution fitting. Then, these values were compared between the before and after period. The average yearly number of crashes and crash rates were also compared to explore the safety impacts of signalization on intersection crashes. Paired t-test was employed to determine if there was a statistically significant difference between the before and after period. In the second phase, statistical crash predictive models were developed to estimate the average number of crashes for all crashes, angle crashes, left-turn crashes, rear-end crashes, and other crashes at intersections before and after the installation of traffic signals. During the modeling process, Poisson regression was performed as the initial step, and negative binomial regression was applied where over-dispersion was tested existing in the crash data. The regression parameters were estimated by using the maximum likelihood method with SAS. The goodness-of-fit for the developed models were evaluated based on Pearson's R-square and likelihood ratio index. In the third phase, case-based crash prediction system was used to predict crash frequencies at a new intersection based on some known cases. In this method, the most similar intersections with respect to roadway environment for application to a new intersection were retrieved from a training database. Then, the information and knowledge from the previous cases were adapted or reused to solve the new case. Subsequently, the predicted crash frequency for the new intersection was evaluated. Once this system was ready, a testing database was used to estimate the number of crashes for intersections with specific characteristics. Lognormal modeling was performed to obtained the final results for this new approach. Finally, the average numbers of crashes at intersections before and after signalization were estimated given the intersection characteristics. The change of the estimated crash frequencies before and after signalization was calculated to represent the impacts of signalization. The estimation of number of crashes and impacts of signalization at intersections were performed based on the statistical modeling approach, operational research modeling approach, and combining both approaches. The results for each one of these approaches are presented in appendices C, D, and E, respectively.

8.2. Conclusions

The following conclusions were made based on crash data analysis, statistical crash prediction modeling and operational research modeling:

- Based on the before and after comparison of mean values, total number of crashes and crash rates increased after signalization. Based on the paired t-test, the following results were statistically significant at a 95% confidence level: total number of crashes increased by 21%, In reference to crash severity, the number of fatal crashes decreased by 13.2% and fatal crash rates decreased by 38% after signal installation. Non-injury crashes increased by 30% for the number of crashes and by 14.8% for crash rates, rear-end crashes had a 102% increase in the number of crashes, and a 47.6% increase in crash rates after signalization. Angle crashes decreased by 14% for number of crashes and by 29% for crash rates. Left turn crashes decreased both in number of crashes and crash rates. Sideswipe crashes increased by 42% in number of crashes. Right turn crash rates decrease by 50.2%. Finally, the following results were not statistically significant at the 95% confidence level: the increase of total crash rates, the decrease after signalization of the number of right turn crashes and crashes related with pedestrian, the increase of crash rates for sideswipe crashes, and the decrease of crash rates for crashes related with pedestrian

- In reference to crash severity, the following results were statistically significant based on the paired t-test at a 95% confidence level: fatal crash rates decreased by 38% after signal installation, injury crashes had an increase of 17.2%, and non-injury number of crashes and non-injury crash rates increased by 30%–and by 14.8%, respectively.. Finally, the following results were not statistically significant at the 95% level of confidence: number of fatal crashes decreased by 13.2%, and injury crash rates decreased by 5% after signalization.

- In regard to surrounding land use, the impact of signalization on intersection crashes was found differently between rural and urban areas. In rural area, the number of crashes decrease was not significant while crash rates decreased significantly. In urban area, the number of crashes and crash rates increased significantly.

- For the statistical modeling, both the "before" and "after" data for all crashes showed extra-variations relative to Poisson distributions. Negative binomial regressions were proved to be appropriate to model the data. For angle and left-turn crashes, the before data showed extra-variation relative to Poisson distribution while the after data did not. For rear-end crashes, Poisson regression was appropriate to model the before data while negative binomial regression was appropriate for the after data. For all other crashes, Poisson regressions were appropriate for both the before and after data.

- Regarding the relative effects of predictor variables in the statistical models and operational models, intersections with higher ADT on the major road would have more crashes than with lower ADT; intersections in urban area would have more crashes than in rural area; intersections located in business area would have more crashes than in other area; intersections with more than four lanes on the major road would have more crashes than those with four or less lanes; intersections with posted speed higher than 45 mph would have less crashes than with posted speed lower or equal to 45 mph; intersections with median on the major road would have more crashes than without median except for rear-end crashes before signalization statistical model; and intersections with paved shoulder would have less crashes than with other types of shoulder.

- Regarding the goodness-of-fit, statistical models developed based on aggregate data (all crashes) performed better than models developed based on disaggregate data (angle, left-turn, rear-end, and other). According to likelihood ratio index, models for all crashes explained more than 20% systematic variation in the crash data while models for angle, left-turn, and other crashes explained less than 10% of the systematic variation in the data. Models for rear-end crashes explained more than 10% systematic variation in the data. In regard to the operational research, lognormal models show very good fit for the data according to the p value

- ADT on the major road was the only predictor variable with estimated parameters significant at the 5% significance level for all developed models.

- In the before and after statistical models for all crashes, the estimated parameters of all seven predictor variables were significant at 20% significance level except for posted speed in the after model. For angle crashes, the estimated parameters of ADT, urban/rural, presence of median, and shoulder treatment were significant for the before model while ADT, land use, number of lanes and presence of median were significant in the after model. For left-turn crashes, the estimated parameters of ADT, land use, number of lanes, presence of median, and shoulder treatment were significant in the before and after models. For rear-end crashes, ADT, land use, number of lanes, posted speed and presence of median are significant in the before model while ADT, number of lanes, posted speed, presence of median and shoulder treatment were significant in the after model. For other crashes, the estimated parameters of ADT, urban/rural, posted speed and median were significant in the before model while ADT, urban/rural, and number of lanes are significant in the after model.

8.3. Recommendations

In future studies, more safety related intersection characteristics are necessary to be considered when doing a before-and after analysis, and/or intersection crash prediction modeling in order to improve the quality of the analysis and the developed models. It is also desirable to have a larger intersection sample. Also, it will be very interesting to develop models for different groups of intersections, such as three-leg or four-leg intersections.

REFERENCES

Agresti, A. (1990). *Categorical Data Analysis*. John Wiley & Sons, New York.

Al-Ghamdi, A., (1991) "Comparison of Accident Rates Using the Likelihood Ratio Testing Technique", Transportation Research Record 1401, pp. 50-54.

Bauer, K. M. and D. W. Harwood (1996). *Statistical Models for At-grade Intersection Accidents.* Report FHWA-RD-96-125. Federal Highway Administration, McLean, VI.

Ben-Akiva, M. and S. R. Lerman (1985). *Discrete Choice Analysis: Theory and Application to Travel Demand*. MIT Press, Cambridge, Mass.

Bernardo, N.R., and J. N. Ivan (1997). "Predicting Number of Crashes Versus Crash Rate Using Poisson Regression", presented at 77th Annual Meeting of the Transportation Research Board.

Bonneson, J. A. and P. T. McCoy (1993). "Estimation of Safety at Two-way Stop-controlled Intersections on Rural Highways". Transportation Research Record, 1401. TRB, Washington, D.C.

Box, P. C. and Associates (1970). *Traffic Control and Roadway Elements - Their Relationship to Highway Safety*. Highway Users Federation for Safety and Mobility.

Datta, T.K., and U. Dutta (1990) "Traffic Signal Installation and Accident Experience", ITE Journal, September pp.39-42

Greene, W. (1997). *Econometric Analysis*. Prentice-Hall, New Jersey.

Hauer, E., J. C. N. Ng, and J.Lovell (1988). "Estimation of Safety at Signalized Intersections". Transportation Research Record 1185, TRB, National Research Council, Washington, D.C.

Huang, Yuanlin (1991) "Accident Prediction Models and Applications for Unsignalized and Signalized Intersections", Proceedings of an international workshop: Intersections without traffic signals II", July, pp. 282-296

Hummer, J. E. (1994). *Traffic Accident Studies. Manual of Transportation Engineering Studies*. Institute of Transportation Engineers. Prentice Hall.

Joshua, S. C. and N. J. Garber (1990). "Estimating Truck Accident Rate and Involvement Using Linear and Poisson Regression Models". Transportation Planning and Technology, Vol. 15.

Jovanis, P. and H. Chang (1985). "Modeling the Relationship of Accidents to Miles Traveled". Transportation Research Record 1068, TRB, National Research Council, Washington, D.C.

King, G. F. and Goldblatt, R. B. (1975), KLD Associates, Inc. “Relationship of Accident Patterns to Type of Intersection Control”. Transportation Research Record, 540. Transportation Research Board. Washing D. C.

Krovvidy, S., and W. G. Wee (1993) “Wastewater-treatment Systems from Case-based Reasoning”, Machine Learning, Vol. 10, 341-363.

Lawless, J. F. (1987). “Negative Binomial and Mixed Poisson Regression”. The Canadian Journal of Statistics, Vol15, No. 3.

McCullagh, P., and J. A. Nelder (1983). *Generalized Linear Models*. Chapman and Hall, London.

Miaou, S., and H. Lum (1993). “Modeling Vehicle Accident and Highway Geometric Design Relationship*s”*. Accidents Analysis and Prevention, Vol. 25, No. 6.

Miaou, S., A. Lu, and H. S. Lum (1996). “Pitfalls of Using R^2 to Evaluate Goodness of Fit of Accident Prediction Models”. Transportation Research Record 1542, TRB, Washington, D.C.

Miaou, S. (1994). “The Relationship between Truck Accidents and Geometric Design of Road Section: Poisson versus Negative Binomial Regression”. Accident Analysis and Prevention. Vol. 26(4).

Nicholson, A.J., (1985) "The Variability of Accident Counts*"*, Accident Analysis & Prediction, Volume. 17. No. 1 pp. 47-56

Ogden, K. W. and S. V. Newstead (1996). “Analysis of Crash Patterns at Signalized Intersections”. Proceedings Roads 96 Conference, Part 5. pp.237-252

Poch, M. and Mannering, F. (1996), “Negative Binomial Analysis of Intersection-Accident Frequencies”, Journal of Transportation Engineering, Vol. 122, No. 2. pp.105-113.

Radwan, A. and D. Wing (1987). *Safety Effects of Traffic Signal Installation. State of Art* Final Report, Arizona Department of Transportation.

Saccomanno, F., and C. Buyco (1988). “Generalized Loglinear Models of Truck Accident Rates”. Paper presented at Transportation Research Board 67th Annual Meeting, Washington, D.C.

Sanchez, M., U. I. R. Cortes, Roda and M. Poch (1997) “Learning and Adaptation in Wastewater Treatment Plants through Case-based Reasoning”, Microcomputers in Civil Engineering, Vol.12, 251-266.

Sayed, T. and F. Rodriguez (1999). “Accident Prediction Models for Urban Unsignalized Intersections in British Columbia”. Transportation Research Record, 1665. TRB, Washington, D. C..

Schank, R. (1982) *Dynamic Memory: A Theory of Learning in Computers and People*, Cambridge University Press, Bambridge, England.

Schank, R., and R. Abelson (1977) *Scripts, Plans, Goals and Understanding*, Lawrence Erlbaum.

Shankar, V., F. Mannering, and W. Barfield (1995). "Effect of Roadway Geometric and Environmental Factors on Rural Accident Frequencies". Accident Analysis and Prevention, Vol. 27, No. 3.

Shen, L.D. (1984) "Before-and-After Accident Study of Newly Signalized Intersections in Washington D.C.", Transportation Planning Journal, Vol. 13 (3), September, pp.425-433

Short, M.S., Woelfl, G.A., Chang, C. (1982). "Effects on Traffic Signal Installation on Accidents", Accident Analysis and Prevention, Vol. 14(2), pp.135-145.

Snyder, D. L. (2000). *Statistical Analysis of the Impacts of Signalization on Intersection Crashes*. Master thesis. Deborah L. Snyder, "*Effects of Intersection Signalization on Crashes",* Master Thesis, University of South Florida, May 2000.

Traffic Safety Facts 1999. National Center for Statistic and Analysis. National Highway Traffic Safety Administration. Washington, D. C..

Vogt, A. and Joe G. Bared, *Accident Models for Two-Lane Rural Roads,* Final Report, Office of Safety and Traffic Operations R&D, Federal Highway Administration.

U.S. Department of Transportation and Federal Highway Administration *Motor Vehicle Accident in Relation to Geometric and Traffic Features of Highway Intersections,* Report No. FHWA-RD-76-128, Volume 1 Executive Summary.

Wettschereck, D., D.W. Aha and T. Mohri (1997). *A review and comparative evaluation of feature weighting methods for lazy learning algorithms*, in *Artificial Intelligence Review*, Kluwer Academic Publishers, New York.

Zhang, J.P. and J. M. Yang (1997). "Instance-based learning for highway accident frequency prediction", *Microcomputers in Civil Engineering*, Vol. 12, pp. 287-294.

APPENDIX A

LIST OF VARIABLES IN FDOT DATABASE

Record Type 00 (Time Log Record)
rectype = Record Type
accnumb = Accident Number
distID = DOT District Number
countyID = DOT County Number
secID = DOT Section Number
subsecID = DOT Subsection Number
milepost = Milepost
year = Accident Year
month = Accident Month
day = Accident Day
counnumb = County Number
citynumb = City Number
node = Node Number
distfrom = Distance from Node
measure = Measurement Type
fromnode = From Node Number
nextnode = Next Node Number
codeable = Codeable Non
typefac = DOT Type Facility
roadtype = DOT Road Type
numblane = DOT Number of Lanes
siteloc = DOT Site Location
skidres = Skid Resistance
friction = Friction Coarse
ADT = ADT
dotnode = DOT Node Number
dotdist = DOT Distance to Node
dotdirec = DOT Direction from Node (Character)
roadnumb = DOT St Road Number
USroad = DOT US Road Number (Character)
rururb = DOT Rural Urban
fedhwy = DOT Fed Hwy Sys
travelw = DOT Travelway
fhptroop =FHP Troop (Character)
fhpdistr = FHP District
accerror = Acc Error Indicator
nodetype = Node Type
fixttype = Fixture Type
sideroad = Side of Road
accsever = Accident Severity
crosstrf = Cross Traffic
classcat = Class Category
milepoin = Node Milepoint A
xtrafveh = Xtraf Veh Miles
accsidrd = Accident Side of Road (character)

acclane = Accident Lane Number
skiddate = Skid Date A
skidnumb = Skid Number
microID = Micro Fish ID

Record Type 01 (Characteristics)
rectype = Record Type
accnumb = Accident Number
distID = DOT District Number
countyID = DOT County Number
secID = DOT Section Number
subsecID = DOT Subsection Number
milepost = Milepost
weekday = Day of the Week
houracc = Hour of Accident
minacc = Minute of Accident
populat =Population
urbrur = DHSMV Urban Rural
sitlocat = DHSMV Site Location
harmev1 = First Harmful Event
harmev2 = Second Harmful Event
offon = Off On Roadway
light = Lighting Conditions
weather = Weather
roadsurf = Road Surface
shoulder = Shoulder Type
surfcond = Road Surface Condition
causerd1 = First Contributing Cause Road
causerd2 = Second Contributing Cause Road
causenv1 = First Contributing Cause Environment
causenv2 = Second Contributing Cause Environment
trafctr1 = First Traffic Control
trafctr2 = Second Traffic Control
trafchar = Trafficway Character
numlanes = DHSMV Number of Lanes
dividnot = DHSMV Divided Not
roadsys = DHSMV Road Sys Indicator
invagent = Investigating Agency
injursev = Injury Severity
damagsev = Damage Severity
insurcod = Insurance Code (Character)
faultcod = Fault Code
alcohol = Alcohol Involved
damagamt = Total Damage Amount
vehdamag = Total Vehicle Damage Amount
propdam = Total Property Damage Amount
totpers = Total Persons

totdriv = Total Drivers
totveh = Total Vehicles
totfatal = Total Fatalities
totnonfa = Total Non Traff Fatals
totinjur = Total Injuries
totpedes = Total Pedestrian
totcycli = Total Pedalcyclist
invagnum = Investigating Agy Number
invcomp = Investigation Complete
hitrun = Hit and Run
locatype = Location Type

Record Type 02 (Vehicle Record)
rectype = Record Type
accnumb = Accident Number
distID = DOT District Number
countyID = DOT County Number
secID = DOT Section Number
subsecID = DOT Subsection Number
milepost = Milepost
formsecn = Form Section Number
vehowndr = Vehicle Owner Driver Same
vehdract = Vehicle Driver Action
vehtype = Vehicle Type
directrv = Direction of Travel (character)
estspeed = Estimated Speed
postsped = Posted Speed
estvehdm = Estimated Vehicle Damage
damtype = Damage Type
pointimp = Point of Impact
vehmov =Vehicle Movement
vehfunc = Vehcile Function
vehcaus1 = First Contributing Cause Vehicle
vehcaus2 = Second Contributing Cause Vehicle
vehrdloc = Vehicle Roadway Location
hazarmat = Hazardous Material
totoccup = Total Occupants
totoccsf = Total Occupants Using Safe
movviol = Moving Violation
vehfault = Vehicle Fault code
vehuse = Vehicle Use
placar = Placarded
oldhazar = Old Hazardous Material

Record Type 03 (Towed Record)
rectype = Record Type
accnumb = Accident Number

distID = DOT District Number
countyID = DOT County Number
secID = DOT Section Number
subsecID = DOT Subsection Number
milepost = Milepost
towtype = Towed Type
towdamag = Towed Damage Amount

Record Type 04 (Driver Record)
rectype = Record Type
accnumb = Accident Number
distID = DOT District Number
countyID = DOT County Number
secID = DOT Section Number
subsecID = DOT Subsection Number
milepost = Milepost
drvage = Driver's Age
drvlictp = Driver License Type
drvbacts = Driver Bac Test
drvbacrs = Driver Bac Results
drvalcol = Driver Alcohol Drug
drvphdef = Driver Physical Defects
drvresid = Driver Residence
drvrace = Driver Race
drvsex = Driver Sex
drvinjsv = Driver Injury Severity
safeeq1 = First Driver Safe Equipment Used
safeeq2 = Second Driver safe Equipment Used
drveject = Driver Eject Code
drvabqst = Driver Ability Question
drvcaus1 = First Contributing Cause Driver
drvcaus2 = Second Contributing Cause Driver
drvcaus3 = Third Contributing Cause Driver
drvofcg1 = First Driver Offense Charged
drvofcg2 = Second Driver Offense Charged
drvofcg3 = Third Driver Offense Charged
reqendor = Required Endorsement

Record Type 05 (Passenger Record)
rectype = Record Type
accnumb = Accident Number
distID = DOT District Number
countyID = DOT County Number
secID = DOT Section Number
subsecID = DOT Subsection Number
milepost = Milepost
pasgage = Passengers Age

pasgloc = Passengers Location
pasgijsv = Passengers Injury Severity
pasgeq1 = First Passengers Safe Equipment
pasgeq2 = Second Passengers Safe Equipment
pasgejct = Passengers Eject Code

Record Type 06 (Pedestrian Record)
rectype = Record Type
accnumb = Accident Number
distID = DOT District Number
countyID = DOT County Number
secID = DOT Section Number
subsecID = DOT Subsection Number
milepost = Milepost
pedage = Pedestrian Age
pedbacts =Pedestrian Bac Test Type
pedbacrs = Pedestrian Bac Test Result
pedalcol = Pedestrian Alcohol Drug
pedphdef = Pedestrian Physical Defect
pedresid = Pedestrian residence
pedrace = Pedestrian Race
pedsex = Pedestrian Sex
pedijsv = Pedestrian Injury Severity
pedcaus1 = First Contributing Cause Pedestrian
pedcaus2 = Second Contributing Cause Pedestrian
pedcaus3 = Third contributing Cause Pedestrian
pedact = Pedestrian Action
pedofcg1 = First Offense Pedestrian Charges
pedofcg2 = Second Offense Pedestrian Charges
pedofcg3 = Third Offense Pedestrian Charges

Record Type 07 (Property Damage Record)
rectype = Record Type
accnumb = Accident Number
distID = DOT District Number
countyID = DOT County Number
secID = DOT Section Number
subsecID = DOT Subsection Number
milepost = Milepost
propdam = Property Damage Amount

Record Type 09 (RCI Features I)
rectype = Record Type
accnumb = Accident Number
distID = DOT District Number
countyID = DOT County Number
secID = DOT Section Number

subsecID = DOT Subsection Number
milepost = Milepost
accesctr =Access Control Type
censuscd = Census Place Code
urbnumb = Urban Area Number
prevland = Prevailing Land use
costcntn = Cost Center Number
statexc = Stationing Exceptions (Character)
assocexc =Associated Station (Character)
widshsh = Width Shoulder to Shoulder
thrsurfw = Thru Surface Width
auxlantp = Auxiliary Lane Type
numauxln = Number Auxiliary Lanes
auxlnw = Auxiliary Lane Width
hwyshtp1 = Hwy Shoulder Type
hwyshtp2 = Hwy Shoulder Type Two
hwyshtp3 = Hwy Shoulder Type Three
hwyshw1 = Hwy Shoulder Width
hwyshw2 = Hwy Shoulder Width Two
hwyshw3 = Hwy Shoulder Width Three
medianw = Median Width
mediantp = Median Type
utstrpw = Utility Strip Width
insshtp1 = Inside Shoulder Type
insshtp2 = Inside Shoulder Type Two
insshtp3 = Inside Shoulder Type Three
insshw1 = Inside Shoulder Width
insshw2 = Inside Shoulder Width Two
insshw3 = Inside Shoulder Width Three
hordeg = Horizontal Degree Curve
horptint = Horizontal PT Intersection
superelv = Super Elevation
percgrad = Percent of Grade
vertcudf = Vertical Curve Deflect
vertptin = Vertical PT Intersection

Record Type 10 (RCI Features II)
rectype = Record Type
accnumb = Accident Number
distID = DOT District Number
countyID = DOT County Number
secID = DOT Section Number
subsecID = DOT Subsection Number
milepost = Milepost
passdist = Passing Sight Distance
rdconsis = Roadway Consistency
rdalign = Roadway Alignment

stopdist = Stopping Sight Distance
pavcond = Pavement Condition
pavindex = Pavement Index
pvsurftp = Pavement Surface Type
pavsurf1 = Pavement Surface 1 (Character)
pavsurf2 = Pavement Surface 2 (Character)
pvlayer1 = Pavement Layer 1
pvlayer2 = Pavement Layer 2
structn = Structure Number
undrpasn = Underpass Number
datspap = Date Speed Approved
datspimp = Date Speed Implemented
maxpstsp = Maximum Posted Speed
minpstsp = Minimum Posted Speed
parkap = Parking Approved
parkimp = Parking Implemented
parklnw = Parking Lane Width
parkrest = Parking Restriction Time
typepark = Type Parking
schsplim = School Speed Limit
tfactor = T Factor
tfacthg = T Factor High
tfactlow = T Factor Low
strpdbwt = Stripes Double White
strpdbyw = Stripes Double Yellow
strpskwb = Stripes Skip Wt Blk
strp skwt = Stripes Skip White
strp skyw = Stripes Skip Yellow
strp sgwt = Stripes Single White
strpsgyw = Stripes Single Yellow

Record Type 11 (RCI Point)
rectype = Record Type
accnumb = Accident Number
distID = DOT District Number
countyID = DOT County Number
secID = DOT Section Number
subsecID = DOT Subsection Number
milepost = Milepost
nodeID = Node ID no City Boundary
reasadj = Reason for Adjustment
lengtadj = Length of Adjustment
unitmeas = Unit of Measurement (Character)
ninsleg = Number of Insect Legs
intercn = Interchange Number
tpinterc = Type of Interchange
rrgradn = National RR Grade Number

trafcnst = Traffic Count Station
datetnap = Date Turn Approved
datetnim = Date Turn Implemented
limtnrst = Limited Turn Restrict
turnrest = Turning Restriction
ngrndsig = Num Ground signs
signcd1 = STD Sign Code One
signcd2 = STD Sign Code Two
signcd3 = STD Sign Code Three
cfmutcd1 = Non Conform MUTCD One1
cfmutcd2 = Non Conform MUTCD Two2
cfmutcd3 = Non Conform MUTCD Three3
sgillum1 = Sign Illumination One
sgillum2 = Sign Illumination Two
sgillum3 = Sign Illumination Three
sgsupor1 = Sign Support One
sgsupor2 = Sign Support Two
sgsupor3 = Sign Support Three
varmesg1 = Variable Message One
varmesg2 = Variable Message Two
varmesg3 = Variable Message Three
nocounsg = Non Counted Signal
trafsgtp = Traffic Signal Type
wrongway = Wrong Way Detector
pointADT = Point ADT
mainbgmp = Main Beginning MP
mainedmp = Main Ending MP
subsectp = Subsection Type
atencond = Attenuator Condition
atenindt = Attenuator Install Date
atenrpdt = Attenuator Repair Date
atenloc = Attenuator Location
atentp = Attenuator Type
genvehdr = General Vehicle Direction
tpinter = Type of Intersection

Record Type 12 (RCI Total)
rectype = Record Type
accnumb = Accident Number
distID = DOT District Number
countyID = DOT County Number
secID = DOT Section Number
subsecID = DOT Subsection Number
milepost = Milepost
natenua = Number of Attenuators
barwallg = Barrier Wall Length
dbgrlg = Double Guardrail Length

miscgrlg = Misc Guardrail Length
stdgrlg = Stand Guardrail Length
nbrgendl = Number Bridge End Delines
ngtdhaz = Number GTD Pst Hzrd Delin
n24ftcrw = Number 24 foot Crosswalks
n36ftcrw = Number 36 foot Crosswalks
n48ftcrw = Number 48 foot Crosswalks
n60ftcrw = Number 60 foot Crosswalks
n72ftcrw = Number 72 foot Crosswalks
n12ftstb = Number 12 foot Stop Bars
n18ftstb = Number 18 foot Stop Bars
n24ftstb = Number 24 foot Stop Bars
n36ftstb = Number 36 foot Stop Bars
n48ftstb = Number 48 foot Stop Bars
nraismrk = Number Raised Markings
ncantliv = Number Cantilever Structs
ngrdpost = Number Ground Post
grdpto30 = Ground Post Over 30
grdptu30 = Ground Post Under 30
noverlan = Number of Overlane Structs

APPENDIX B

PRELIMINARY SELECTED VARIABLES

NO.	VARIABLE	DESCRIPTION
1	ACCNUMB	ACCIDENT NUMBER
2	DISTID	DOT DISTRICT NUMBER
3	COUNTYID	DOT COUNTY NUMBER
4	SECID	DOT SECTION NUMBER
5	SUBSECID	DOT SUBSECTION NUMBER
6	MILEPOST	MILEPOST OF ACCIDENT SITE
7	YEAR	ACCIDENT YEAR
8	ROADTYPE	DOT ROAD TYPE
9	NUMBLANE	DOT NUMBER OF LANE
10	SITELOC	DOT SITE LOCATION
11	ADT	ADT OF MAJOR ROAD
12	RURURB	DOT RURAL/URBAN
13	FEDHWY	DOT FED HWY SYS
14	TRAVELW	DOT TRAVEL WAY
15	ACCSEVER	ACCIDENT SEVERITY
16	CROSSTRF	CROSS TRAFFIC
17	CLASSCAT	CLASS CATEGORY
18	URBRUR	DHSMV RURAL/URBAN
19	SITLOCAT	DHSMV SITE LOCATION
20	HARMEV1	FIRST HARMFUL EVENT
21	ROADSURF	ROAD SURFACE
22	SHOULDER	SHOULDER TREATMENT TYPE
23	TRAFCTR1	FIRST TRAFFIC CONTROL
24	TRAFCHAR	TRAFFIC WAY CHRACTER
25	NUMLANES	DHSMV NUMBER OF LANE
26	DIVIDNOT	DHSMV DIVIDED/NOT
27	ROADSYS	ROADWAY SYSTEM INDICATOR
28	INJURSEV	INJURY SEVERITY
29	DAMAGSEV	DAMAGE SEVERITY
30	ALCOHOL	ALCOHOL INVOLVED/NOT
31	DAMAGAMT	TOTAL DAMAGE AMOUNT
32	VEHDAMAG	TOTAL VEHICLE DAMAGE AMOUNT
33	PROPDAM	TOTAL PROPERTY DAMAGE
34	TOTFATAL	TOTAL FATALITIES
35	TOTINJUR	TOTAL INJURIES

Preliminary Selected Variables (Continued)

NO.	VARIABLE	DESCRIPTION
36	LOCATYPE	LOCATION TYPE
37	VEHTYPE	VVEHICLE TYPE
38	ESTSPEED	ESTIMATED SPEED
39	POSTSPED	POSTED SPEED
40	POINTIMP	POINT OF IMPACT
41	ACCESCTR	ACCESS CONTROL TYPE
42	PREVLAND	PREVAILING LAND USE
43	HWYSHTP1	HIGHWAY SHOULDER TYPE
44	HWYSHW1	HIGHWAY SHOULDER WIDTH
45	MEDIANW	MEDIAN WIDTH
46	MEDIANTP	MEDIAN TYPE
47	HORPTINT	HORIZONTAL PT INTERSECTION
48	SUPERELV	SUPER ELEVATION
49	VERTPTIN	VERTICAL PT INTERSECTION
50	PASSDIST	PASSING DISTANCE
51	RDCONSIS	ROAD CONSISTENCY
52	RDALIGN	ROADWAY ALIGNMENT
53	STOPDIST	STOPPING DISTANCE
54	PAVCOND	PAVEMENT CONDITION
55	PAVINDEX	PAVEMENT INDEX
56	PVSURFTP	PAVEMENT SURFACE TYPE
57	MAXPSTSP	MAXIMUM POSTED SPEED
58	MINPSTSP	MINIMUM POSTED SPEED
59	NINSLEG	NUMBER OF INTERSECTION LEGS
60	TRAFSGTP	TRAFFIC SIGNAL TYPE
61	POINTADT	POINT ADT
62	TPINTER	TYPE OF INTERSECTION

APPENDIX C

STATISTICAL MODELING RESULTS

Total Crashes – Table One
Statistical Modeling

Surrounding Land Use: Urban   Location Type: Business   Shoulder: Paved

Intersection Characteristics				Total Number of Crashes			
Speed	Lanes	Median	ADT (vpd)	Before	After	Increase	% Increase
≤ 45 mph	≤ 4	Divided	≤ 15000	2.94	2.49	-0.45	-15.30
			15000 to ≤ 30000	3.89	4.06	0.17	4.40
			> 30000	5.13	6.60	1.47	28.69
		Undivided	≤ 15000	2.24	2.07	-0.16	-7.25
			15000 to ≤ 30000	2.95	3.37	0.42	14.32
			> 30000	3.90	5.49	1.59	40.90
	> 4	Divided	≤ 15000	3.82	3.25	-0.57	-14.96
			15000 to ≤ 30000	5.05	5.29	0.24	4.82
			> 30000	6.66	8.61	1.95	29.20
		Undivided	≤ 15000	2.90	2.70	-0.20	-6.88
			15000 to ≤ 30000	3.83	4.40	0.57	14.77
			> 30000	5.06	7.16	2.10	41.47
> 45 mph	≤ 4	Divided	≤ 15000	2.49	2.49	0.01	0.35
			15000 to ≤ 30000	3.28	4.06	0.78	23.69
			> 30000	4.33	6.60	2.27	52.46
		Undivided	≤ 15000	1.89	2.07	0.19	9.88
			15000 to ≤ 30000	2.49	3.37	0.88	35.43
			> 30000	3.29	5.49	2.20	66.93
	> 4	Divided	≤ 15000	3.23	3.25	0.02	0.75
			15000 to ≤ 30000	4.26	5.29	1.03	24.18
			> 30000	5.62	8.61	2.98	53.07
		Undivided	≤ 15000	2.45	2.70	0.25	10.32
			15000 to ≤ 30000	3.24	4.40	1.16	35.97
			> 30000	4.27	7.16	2.89	67.60

Note: Increase = After – Before

Total Crashes – Table Two
Statistical Modeling

Surrounding Land Use: Urban   Location Type: Business   Shoulder: Other

Intersection Characteristics				Total Number of Crashes			
Speed	Lanes	Median	ADT (vpd)	Before	After	Increase	% Increase
≤ 45 mph	≤ 4	Divided	≤ 15000	3.48	2.78	-0.70	-20.04
			15000 to ≤ 30000	4.60	4.53	-0.07	-1.45
			> 30000	6.07	7.37	1.30	21.47
		Undivided	≤ 15000	2.64	2.32	-0.33	-12.45
			15000 to ≤ 30000	3.49	3.77	0.28	7.91
			> 30000	4.61	6.13	1.52	33.00
	> 4	Divided	≤ 15000	4.52	3.63	-0.89	-19.72
			15000 to ≤ 30000	5.97	5.91	-0.06	-1.05
			> 30000	7.88	9.61	1.73	21.96
		Undivided	≤ 15000	3.43	3.02	-0.42	-12.10
			15000 to ≤ 30000	4.53	4.91	0.38	8.34
			> 30000	5.99	7.99	2.01	33.54
> 45 mph	≤ 4	Divided	≤ 15000	2.94	2.78	-0.16	-5.28
			15000 to ≤ 30000	3.88	4.53	0.65	16.75
			> 30000	5.12	7.37	2.25	43.91
		Undivided	≤ 15000	2.23	2.32	0.08	3.72
			15000 to ≤ 30000	2.95	3.77	0.82	27.84
			> 30000	3.89	6.13	2.24	57.57
	> 4	Divided	≤ 15000	3.82	3.63	-0.19	-4.90
			15000 to ≤ 30000	5.04	5.91	0.87	17.22
			> 30000	6.65	9.61	2.96	44.48
		Undivided	≤ 15000	2.90	3.02	0.12	4.13
			15000 to ≤ 30000	3.83	4.91	1.09	28.35
			> 30000	5.05	7.99	2.94	58.20

Note: Increase = After – Before

Total Crashes – Table Three
Statistical Modeling

Surrounding Land Use: Urban Location Type: Other Shoulder: Paved

Intersection Characteristics				Total Number of Crashes			
Speed	Lanes	Median	ADT (vpd)	Before	After	Increase	% Increase
≤ 45 mph	≤ 4	Divided	≤ 15000	2.48	2.10	-0.38	-15.49
			15000 to ≤ 30000	3.28	3.41	0.14	4.16
			> 30000	4.33	5.55	1.23	28.39
		Undivided	≤ 15000	1.89	1.74	-0.14	-7.47
			15000 to ≤ 30000	2.49	2.84	0.35	14.05
			> 30000	3.29	4.62	1.33	40.58
	> 4	Divided	≤ 15000	3.22	2.74	-0.49	-15.15
			15000 to ≤ 30000	4.26	4.45	0.20	4.58
			> 30000	5.62	7.24	1.62	28.90
		Undivided	≤ 15000	2.45	2.27	-0.17	-7.10
			15000 to ≤ 30000	3.23	3.70	0.47	14.51
			> 30000	4.27	6.02	1.76	41.14
> 45 mph	≤ 4	Divided	≤ 15000	2.10	2.10	0.00	0.12
			15000 to ≤ 30000	2.77	3.41	0.65	23.40
			> 30000	3.65	5.55	1.90	52.10
		Undivided	≤ 15000	1.59	1.74	0.15	9.63
			15000 to ≤ 30000	2.10	2.84	0.74	35.12
			> 30000	2.77	4.62	1.85	66.55
	> 4	Divided	≤ 15000	2.72	2.74	0.01	0.52
			15000 to ≤ 30000	3.59	4.45	0.86	23.90
			> 30000	4.74	7.24	2.50	52.71
		Undivided	≤ 15000	2.07	2.27	0.21	10.06
			15000 to ≤ 30000	2.73	3.70	0.97	35.66
			> 30000	3.60	6.02	2.42	67.21

Note: Increase = After – Before

Total Crashes – Table Four
Statistical Modeling

Surrounding Land Use: Urban   Location Type: Other   Shoulder: Other

Intersection Characteristics				Total Number of Crashes			
Speed	Lanes	Median	ADT (vpd)	Before	After	Increase	% Increase
≤ 45 mph	≤ 4	Divided	≤ 15000	2.94	2.34	-0.59	-20.23
			15000 to ≤ 30000	3.88	3.81	-0.06	-1.68
			> 30000	5.12	6.20	1.08	21.19
		Undivided	≤ 15000	2.23	1.95	-0.28	-12.65
			15000 to ≤ 30000	2.94	3.17	0.23	7.66
			> 30000	3.89	5.16	1.27	32.70
	> 4	Divided	≤ 15000	3.81	3.05	-0.76	-19.91
			15000 to ≤ 30000	5.03	4.97	-0.06	-1.28
			> 30000	6.65	8.09	1.44	21.68
		Undivided	≤ 15000	2.90	2.54	-0.36	-12.30
			15000 to ≤ 30000	3.82	4.13	0.31	8.09
			> 30000	5.05	6.72	1.68	33.23
> 45 mph	≤ 4	Divided	≤ 15000	2.48	2.34	-0.14	-5.49
			15000 to ≤ 30000	3.27	3.81	0.54	16.49
			> 30000	4.32	6.20	1.88	43.58
		Undivided	≤ 15000	1.88	1.95	0.07	3.48
			15000 to ≤ 30000	2.48	3.17	0.68	27.55
			> 30000	3.28	5.16	1.88	57.21
	> 4	Divided	≤ 15000	3.22	3.05	-0.16	-5.11
			15000 to ≤ 30000	4.25	4.97	0.72	16.95
			> 30000	5.61	8.09	2.48	44.15
		Undivided	≤ 15000	2.44	2.54	0.10	3.89
			15000 to ≤ 30000	3.23	4.13	0.91	28.06
			> 30000	4.26	6.72	2.46	57.84

Note: Increase = After – Before

Total Crashes – Table Five
Statistical Modeling

Surrounding Land Use: Rural   Location Type: Business   Shoulder: Paved

Intersection Characteristics				Total Number of Crashes			
Speed	Lanes	Median	ADT (vpd)	Before	After	Increase	% Increase
≤ 45 mph	≤ 4	Divided	≤ 15000	2.61	2.27	-0.34	-13.20
			15000 to ≤ 30000	3.45	3.69	0.24	6.98
			> 30000	4.55	6.00	1.45	31.86
		Undivided	≤ 15000	1.98	1.89	-0.10	-4.96
			15000 to ≤ 30000	2.62	3.07	0.45	17.14
			> 30000	3.46	4.99	1.53	44.38
	> 4	Divided	≤ 15000	3.39	2.96	-0.44	-12.86
			15000 to ≤ 30000	4.48	4.81	0.33	7.41
			> 30000	5.91	7.83	1.92	32.39
		Undivided	≤ 15000	2.58	2.46	-0.12	-4.58
			15000 to ≤ 30000	3.40	4.00	0.60	17.61
			> 30000	4.49	6.51	2.02	44.96
> 45 mph	≤ 4	Divided	≤ 15000	2.21	2.27	0.06	2.83
			15000 to ≤ 30000	2.91	3.69	0.78	26.74
			> 30000	3.84	6.00	2.16	56.22
		Undivided	≤ 15000	1.67	1.89	0.21	12.59
			15000 to ≤ 30000	2.21	3.07	0.86	38.78
			> 30000	2.92	4.99	2.07	71.05
	> 4	Divided	≤ 15000	2.86	2.96	0.09	3.24
			15000 to ≤ 30000	3.78	4.81	1.03	27.25
			> 30000	4.99	7.83	2.84	56.85
		Undivided	≤ 15000	2.18	2.46	0.28	13.04
			15000 to ≤ 30000	2.87	4.00	1.13	39.33
			> 30000	3.79	6.51	2.72	71.74

Note: Increase = After – Before

Total Crashes – Table Six
Statistical Modeling

Surrounding Land Use: Rural   Location Type: Business   Shoulder: Other

Intersection Characteristics				Total Number of Crashes			
Speed	Lanes	Median	ADT (vpd)	Before	After	Increase	% Increase
≤ 45 mph	≤ 4	Divided	≤ 15000	3.09	2.53	-0.56	-18.07
			15000 to ≤ 30000	4.08	4.12	0.04	0.98
			> 30000	5.39	6.70	1.32	24.47
		Undivided	≤ 15000	2.35	2.11	-0.24	-10.29
			15000 to ≤ 30000	3.10	3.43	0.33	10.57
			> 30000	4.09	5.57	1.48	36.29
	> 4	Divided	≤ 15000	4.01	3.30	-0.71	-17.74
			15000 to ≤ 30000	5.30	5.37	0.07	1.39
			> 30000	7.00	8.74	1.75	24.97
		Undivided	≤ 15000	3.05	2.75	-0.30	-9.93
			15000 to ≤ 30000	4.02	4.47	0.44	11.02
			> 30000	5.31	7.27	1.96	36.83
> 45 mph	≤ 4	Divided	≤ 15000	2.61	2.53	-0.08	-2.94
			15000 to ≤ 30000	3.44	4.12	0.68	19.64
			> 30000	4.55	6.70	2.16	47.46
		Undivided	≤ 15000	1.98	2.11	0.12	6.28
			15000 to ≤ 30000	2.62	3.43	0.81	31.00
			> 30000	3.45	5.57	2.12	61.46
	> 4	Divided	≤ 15000	3.39	3.30	-0.09	-2.55
			15000 to ≤ 30000	4.47	5.37	0.90	20.12
			> 30000	5.90	8.74	2.84	48.05
		Undivided	≤ 15000	2.57	2.75	0.17	6.71
			15000 to ≤ 30000	3.40	4.47	1.07	31.52
			> 30000	4.48	7.27	2.78	62.11

Note: Increase = After – Before

Total Crashes – Table Seven
Statistical Modeling

Surrounding Land Use: Rural　　Location Type: Other　　Shoulder: Paved

Intersection Characteristics				Total Number of Crashes			
Speed	Lanes	Median	ADT (vpd)	Before	After	Increase	% Increase
≤ 45 mph	≤ 4	Divided	≤ 15000	2.20	1.91	-0.30	-13.40
			15000 to ≤ 30000	2.91	3.10	0.20	6.74
			> 30000	3.84	5.05	1.21	31.56
		Undivided	≤ 15000	1.67	1.59	-0.09	-5.18
			15000 to ≤ 30000	2.21	2.58	0.37	16.87
			> 30000	2.92	4.20	1.28	44.05
	> 4	Divided	≤ 15000	2.86	2.49	-0.37	-13.06
			15000 to ≤ 30000	3.78	4.05	0.27	7.17
			> 30000	4.99	6.59	1.60	32.09
		Undivided	≤ 15000	2.17	2.07	-0.10	-4.80
			15000 to ≤ 30000	2.87	3.37	0.50	17.34
			> 30000	3.79	5.48	1.69	44.63
> 45 mph	≤ 4	Divided	≤ 15000	1.86	1.91	0.05	2.59
			15000 to ≤ 30000	2.46	3.10	0.65	26.45
			> 30000	3.24	5.05	1.81	55.86
		Undivided	≤ 15000	1.41	1.59	0.17	12.33
			15000 to ≤ 30000	1.86	2.58	0.72	38.46
			> 30000	2.46	4.20	1.74	70.66
	> 4	Divided	≤ 15000	2.42	2.49	0.07	3.00
			15000 to ≤ 30000	3.19	4.05	0.86	26.96
			> 30000	4.21	6.59	2.38	56.49
		Undivided	≤ 15000	1.83	2.07	0.23	12.78
			15000 to ≤ 30000	2.42	3.37	0.94	39.01
			> 30000	3.20	5.48	2.28	71.34

Note: Increase = After – Before

Total Crashes – Table Eight
Statistical Modeling

Surrounding Land Use: Rural | Location Type: Other | Shoulder: Other

Intersection Characteristics				Total Number of Crashes			
Speed	Lanes	Median	ADT (vpd)	Before	After	Increase	% Increase
≤ 45 mph	≤ 4	Divided	≤ 15000	2.61	2.13	-0.48	-18.26
			15000 to ≤ 30000	3.44	3.47	0.03	0.75
			> 30000	4.54	5.64	1.10	24.18
		Undivided	≤ 15000	1.98	1.77	-0.21	-10.50
			15000 to ≤ 30000	2.61	2.88	0.27	10.32
			> 30000	3.45	4.69	1.24	35.97
	> 4	Divided	≤ 15000	3.38	2.78	-0.61	-17.93
			15000 to ≤ 30000	4.47	4.52	0.05	1.16
			> 30000	5.90	7.35	1.46	24.68
		Undivided	≤ 15000	2.57	2.31	-0.26	-10.14
			15000 to ≤ 30000	3.39	3.76	0.37	10.76
			> 30000	4.48	6.12	1.64	36.52
> 45 mph	≤ 4	Divided	≤ 15000	2.20	2.13	-0.07	-3.16
			15000 to ≤ 30000	2.90	3.47	0.56	19.36
			> 30000	3.83	5.64	1.81	47.12
		Undivided	≤ 15000	1.67	1.77	0.10	6.04
			15000 to ≤ 30000	2.21	2.88	0.68	30.70
			> 30000	2.91	4.69	1.78	61.09
	> 4	Divided	≤ 15000	2.86	2.78	-0.08	-2.77
			15000 to ≤ 30000	3.77	4.52	0.75	19.84
			> 30000	4.98	7.35	2.38	47.71
		Undivided	≤ 15000	2.17	2.31	0.14	6.46
			15000 to ≤ 30000	2.86	3.76	0.89	31.22
			> 30000	3.78	6.12	2.33	61.74

Note: Increase = After – Before

Angle Crashes – Table One
Statistical Modeling

Surrounding Land Use: Urban Location Type: Business Shoulder: Paved
Speed: ≤ 45 mph or > 45 mph

Intersection Characteristics			Number of Angle Crashes			
Lanes	Median	ADT (vpd)	Before	After	Increase	% Increase
≤ 4	Divided	≤ 15000	0.77	0.59	-0.18	-22.96
		15000 to ≤ 30000	0.96	0.80	-0.17	-17.49
		> 30000	1.21	1.07	-0.14	-11.62
	Undivided	≤ 15000	0.51	0.50	-0.01	-1.23
		15000 to ≤ 30000	0.63	0.67	0.04	5.79
		> 30000	0.79	0.90	0.11	13.31
> 4	Divided	≤ 15000	0.77	0.86	0.08	11.02
		15000 to ≤ 30000	0.96	1.15	0.18	18.91
		> 30000	1.21	1.54	0.33	27.37
	Undivided	≤ 15000	0.51	0.72	0.21	42.33
		15000 to ≤ 30000	0.63	0.97	0.33	52.46
		> 30000	0.79	1.30	0.50	63.30

Note: Increase = After – Before

Angle Crashes – Table Two
Statistical Modeling

Surrounding Land Use: Urban Location Type: Business Shoulder: Other
Speed: ≤ 45 mph or > 45 mph

Intersection Characteristics			Number of Angle Crashes			
Lanes	Median	ADT (vpd)	Before	After	Increase	% Increase
≤ 4	Divided	≤ 15000	1.15	0.59	-0.55	-48.28
		15000 to ≤ 30000	1.44	0.80	-0.64	-44.61
		> 30000	1.80	1.07	-0.73	-40.67
	Undivided	≤ 15000	0.76	0.50	-0.25	-33.69
		15000 to ≤ 30000	0.95	0.67	-0.27	-28.98
		> 30000	1.18	0.90	-0.28	-23.93
> 4	Divided	≤ 15000	1.15	0.86	-0.29	-25.47
		15000 to ≤ 30000	1.44	1.15	-0.29	-20.17
		> 30000	1.80	1.54	-0.26	-14.50
	Undivided	≤ 15000	0.76	0.72	-0.03	-4.45
		15000 to ≤ 30000	0.95	0.97	0.02	2.35
		> 30000	1.18	1.30	0.11	9.63

Note: Increase = After – Before

Angle Crashes – Table Three
Statistical Modeling

Surrounding Land Use: Urban   Location Type: Other   Shoulder: Paved
Speed: ≤ 45 mph or > 45 mph

Intersection Characteristics			Number of Angle Crashes			
Lanes	Median	ADT (vpd)	Before	After	Increase	% Increase
≤ 4	Divided	≤ 15000	0.77	0.50	-0.27	-35.45
		15000 to ≤ 30000	0.96	0.67	-0.30	-30.86
		> 30000	1.21	0.89	-0.31	-25.94
	Undivided	≤ 15000	0.51	0.42	-0.09	-17.24
		15000 to ≤ 30000	0.63	0.56	-0.07	-11.35
		> 30000	0.79	0.75	-0.04	-5.05
> 4	Divided	≤ 15000	0.77	0.72	-0.05	-6.97
		15000 to ≤ 30000	0.96	0.96	0.00	-0.36
		> 30000	1.21	1.29	0.08	6.73
	Undivided	≤ 15000	0.51	0.61	0.10	19.27
		15000 to ≤ 30000	0.63	0.81	0.18	27.75
		> 30000	0.79	1.09	0.29	36.83

Note: Increase = After – Before

Angle Crashes – Table Four
Statistical Modeling

Surrounding Land Use: Urban Location Type: Other Shoulder: Other
Speed: ≤ 45 mph or > 45 mph

Intersection Characteristics			Number of Angle Crashes			
Lanes	Median	ADT (vpd)	Before	After	Increase	% Increase
≤ 4	Divided	≤ 15000	1.15	0.50	-0.65	-56.66
		15000 to ≤ 30000	1.44	0.67	-0.77	-53.58
		> 30000	1.80	0.89	-0.90	-50.28
	Undivided	≤ 15000	0.76	0.42	-0.34	-44.44
		15000 to ≤ 30000	0.95	0.56	-0.38	-40.49
		> 30000	1.18	0.75	-0.43	-36.26
> 4	Divided	≤ 15000	1.15	0.72	-0.43	-37.55
		15000 to ≤ 30000	1.44	0.96	-0.48	-33.11
		> 30000	1.80	1.29	-0.51	-28.35
	Undivided	≤ 15000	0.76	0.61	-0.15	-19.93
		15000 to ≤ 30000	0.95	0.81	-0.13	-14.24
		> 30000	1.18	1.09	-0.10	-8.14

Note: Increase = After – Before

Angle Crashes – Table Five
Statistical Modeling

Surrounding Land Use: Rural   Location Type: Business   Shoulder: Paved
Speed: ≤ 45 mph or > 45 mph

Intersection Characteristics			Number of Angle Crashes			
Lanes	Median	ADT (vpd)	Before	After	Increase	% Increase
≤ 4	Divided	≤ 15000	0.58	0.59	0.01	2.22
		15000 to ≤ 30000	0.73	0.80	0.07	9.49
		> 30000	0.91	1.07	0.16	17.28
	Undivided	≤ 15000	0.38	0.50	0.12	31.06
		15000 to ≤ 30000	0.48	0.67	0.19	40.38
		> 30000	0.60	0.90	0.30	50.37
> 4	Divided	≤ 15000	0.58	0.86	0.27	47.31
		15000 to ≤ 30000	0.73	1.15	0.42	57.79
		> 30000	0.91	1.54	0.63	69.01
	Undivided	≤ 15000	0.38	0.72	0.34	88.87
		15000 to ≤ 30000	0.48	0.97	0.49	102.30
		> 30000	0.60	1.30	0.70	116.69

Note: Increase = After – Before

Angle Crashes – Table Six
Statistical Modeling

Surrounding Land Use: Rural   Location Type: Business   Shoulder: Other
Speed: ≤ 45 mph or > 45 mph

Intersection Characteristics			Number of Angle Crashes			
Lanes	Median	ADT (vpd)	Before	After	Increase	% Increase
≤ 4	Divided	≤ 15000	0.87	0.59	-0.27	-31.37
≤ 4	Divided	15000 to ≤ 30000	1.08	0.80	-0.29	-26.49
≤ 4	Divided	> 30000	1.35	1.07	-0.29	-21.27
≤ 4	Undivided	≤ 15000	0.57	0.50	-0.07	-12.01
≤ 4	Undivided	15000 to ≤ 30000	0.71	0.67	-0.04	-5.76
≤ 4	Undivided	> 30000	0.89	0.90	0.01	0.94
> 4	Divided	≤ 15000	0.87	0.86	-0.01	-1.10
> 4	Divided	15000 to ≤ 30000	1.08	1.15	0.06	5.93
> 4	Divided	> 30000	1.35	1.54	0.18	13.46
> 4	Undivided	≤ 15000	0.57	0.72	0.15	26.79
> 4	Undivided	15000 to ≤ 30000	0.71	0.97	0.26	35.81
> 4	Undivided	> 30000	0.89	1.30	0.41	45.47

Note: Increase = After – Before

Angle Crashes – Table Seven
Statistical Modeling

Surrounding Land Use: Rural   Location Type: Other   Shoulder: Paved
Speed: ≤ 45 mph or > 45 mph

Intersection Characteristics			Number of Angle Crashes			
Lanes	Median	ADT (vpd)	Before	After	Increase	% Increase
≤ 4	Divided	≤ 15000	0.58	0.50	-0.08	-14.34
		15000 to ≤ 30000	0.73	0.67	-0.06	-8.25
		> 30000	0.91	0.89	-0.02	-1.72
	Undivided	≤ 15000	0.38	0.42	0.04	9.82
		15000 to ≤ 30000	0.48	0.56	0.08	17.63
		> 30000	0.60	0.75	0.16	26.00
> 4	Divided	≤ 15000	0.58	0.72	0.14	23.44
		15000 to ≤ 30000	0.73	0.96	0.23	32.22
		> 30000	0.91	1.29	0.38	41.62
	Undivided	≤ 15000	0.38	0.61	0.22	58.26
		15000 to ≤ 30000	0.48	0.81	0.33	69.52
		> 30000	0.60	1.09	0.49	81.58

Note: Increase = After – Before

Angle Crashes – Table Eight
Statistical Modeling

Surrounding Land Use: Rural   Location Type: Other   Shoulder: Other
Speed: ≤ 45 mph or > 45 mph

Intersection Characteristics			Number of Angle Crashes			
Lanes	Median	ADT (vpd)	Before	After	Increase	% Increase
≤ 4	Divided	≤ 15000	0.87	0.50	-0.37	-42.50
		15000 to ≤ 30000	1.08	0.67	-0.42	-38.41
		> 30000	1.35	0.89	-0.46	-34.03
	Undivided	≤ 15000	0.57	0.42	-0.15	-26.27
		15000 to ≤ 30000	0.71	0.56	-0.15	-21.03
		> 30000	0.89	0.75	-0.14	-15.41
> 4	Divided	≤ 15000	0.87	0.72	-0.15	-17.13
		15000 to ≤ 30000	1.08	0.96	-0.12	-11.24
		> 30000	1.35	1.29	-0.07	-4.92
	Undivided	≤ 15000	0.57	0.61	0.04	6.25
		15000 to ≤ 30000	0.71	0.81	0.10	13.80
		> 30000	0.89	1.09	0.20	21.90

Note: Increase = After – Before

Left-turn Crashes – Table One
Statistical Modeling

Surrounding Land Use: Rural or Urban | Location Type: Business
Shoulder: Paved | Speed: ≤ 45 mph or > 45 mph

Intersection Characteristics			Number of Left Turn Crashes			
Lanes	Median	ADT (vpd)	Before	After	Increase	% Increase
≤ 4	Divided	≤ 15000	0.61	0.59	-0.02	-2.91
		15000 to ≤ 30000	0.77	0.71	-0.06	-7.70
		> 30000	0.97	0.85	-0.12	-12.25
	Undivided	≤ 15000	0.49	0.35	-0.14	-28.13
		15000 to ≤ 30000	0.62	0.42	-0.20	-31.68
		> 30000	0.78	0.51	-0.27	-35.05
> 4	Divided	≤ 15000	0.85	0.79	-0.06	-7.57
		15000 to ≤ 30000	1.07	0.94	-0.13	-12.13
		> 30000	1.35	1.13	-0.22	-16.46
	Undivided	≤ 15000	0.69	0.47	-0.22	-31.58
		15000 to ≤ 30000	0.86	0.56	-0.30	-34.96
		> 30000	1.09	0.67	-0.41	-38.16

Note: Increase = After – Before

Left-turn Crashes – Table Two
Statistical Modeling

Surrounding Land Use: Rural or Urban — Location Type: Business
Shoulder: Other — Speed: ≤ 45 mph or > 45 mph

Intersection Characteristics			Number of Left Turn Crashes			
Lanes	Median	ADT (vpd)	Before	After	Increase	% Increase
≤ 4	Divided	≤ 15000	0.79	0.70	-0.08	-10.55
		15000 to ≤ 30000	0.99	0.84	-0.15	-14.96
		> 30000	1.24	1.00	-0.24	-19.16
	Undivided	≤ 15000	0.63	0.42	-0.21	-33.79
		15000 to ≤ 30000	0.80	0.50	-0.30	-37.05
		> 30000	1.00	0.60	-0.40	-40.16
> 4	Divided	≤ 15000	1.09	0.93	-0.16	-14.85
		15000 to ≤ 30000	1.37	1.11	-0.26	-19.05
		> 30000	1.73	1.33	-0.40	-23.04
	Undivided	≤ 15000	0.88	0.56	-0.33	-36.97
		15000 to ≤ 30000	1.11	0.66	-0.44	-40.08
		> 30000	1.39	0.79	-0.60	-43.03

Note: Increase = After – Before

Left-turn Crashes – Table Three
Statistical Modeling

Surrounding Land Use: Rural or Urban   Location Type: Other
Shoulder: Paved   Speed: ≤ 45 mph or > 45 mph

Intersection Characteristics			Number of Left Turn Crashes			
Lanes	Median	ADT (vpd)	Before	After	Increase	% Increase
≤ 4	Divided	≤ 15000	0.44	0.47	0.03	6.55
		15000 to ≤ 30000	0.56	0.57	0.01	1.29
		> 30000	0.70	0.68	-0.03	-3.71
	Undivided	≤ 15000	0.36	0.28	-0.08	-21.13
		15000 to ≤ 30000	0.45	0.34	-0.11	-25.02
		> 30000	0.57	0.40	-0.16	-28.72
> 4	Divided	≤ 15000	0.62	0.63	0.01	1.43
		15000 to ≤ 30000	0.78	0.75	-0.03	-3.57
		> 30000	0.98	0.90	-0.08	-8.33
	Undivided	≤ 15000	0.50	0.37	-0.12	-24.92
		15000 to ≤ 30000	0.63	0.45	-0.18	-28.62
		> 30000	0.79	0.53	-0.25	-32.15

Note: Increase = After – Before

Left-turn Crashes – Table Four
Statistical Modeling

Surrounding Land Use: Rural or Urban — Location Type: Other
Shoulder: Other — Speed: ≤ 45 mph or > 45 mph

Intersection Characteristics			Number of Left Turn Crashes			
Lanes	Median	ADT (vpd)	Before	After	Increase	% Increase
≤ 4	Divided	≤ 15000	0.57	0.56	-0.01	-1.84
		15000 to ≤ 30000	0.72	0.67	-0.05	-6.69
		> 30000	0.90	0.80	-0.10	-11.29
	Undivided	≤ 15000	0.46	0.33	-0.13	-27.34
		15000 to ≤ 30000	0.58	0.40	-0.18	-30.93
		> 30000	0.73	0.48	-0.25	-34.33
> 4	Divided	≤ 15000	0.79	0.74	-0.05	-6.56
		15000 to ≤ 30000	1.00	0.89	-0.11	-11.17
		> 30000	1.25	1.06	-0.19	-15.55
	Undivided	≤ 15000	0.64	0.44	-0.20	-30.83
		15000 to ≤ 30000	0.80	0.53	-0.28	-34.24
		> 30000	1.01	0.63	-0.38	-37.49

Note: Increase = After – Before

Rear-end Crashes – Table One
Statistical Modeling

Surrounding Land Use: Rural or Urban — Location Type: Business
Shoulder: Paved

Speed	Lanes	Median	ADT (vpd)	Before	After	Increase	% Increase
≤ 45 mph	≤ 4	Divided	≤ 15000	0.45	0.69	0.24	52.29
			15000 to ≤ 30000	0.79	1.32	0.53	66.31
			> 30000	1.38	2.51	1.13	81.63
		Undivided	≤ 15000	0.53	0.55	0.02	4.25
			15000 to ≤ 30000	0.93	1.05	0.13	13.85
			> 30000	1.61	2.01	0.39	24.33
	> 4	Divided	≤ 15000	0.69	0.84	0.15	21.81
			15000 to ≤ 30000	1.21	1.60	0.40	33.03
			> 30000	2.10	3.06	0.95	45.28
		Undivided	≤ 15000	0.81	0.67	-0.13	-16.61
			15000 to ≤ 30000	1.41	1.28	-0.13	-8.94
			> 30000	2.46	2.45	-0.01	-0.55
> 45 mph	≤ 4	Divided	≤ 15000	0.28	0.60	0.31	110.75
			15000 to ≤ 30000	0.49	1.14	0.64	130.16
			> 30000	0.86	2.17	1.31	151.36
		Undivided	≤ 15000	0.33	0.48	0.15	44.27
			15000 to ≤ 30000	0.58	0.91	0.33	57.55
			> 30000	1.01	1.74	0.73	72.06
	> 4	Divided	≤ 15000	0.43	0.73	0.30	68.57
			15000 to ≤ 30000	0.75	1.39	0.63	84.10
			> 30000	1.31	2.64	1.33	101.05
		Undivided	≤ 15000	0.50	0.58	0.08	15.40
			15000 to ≤ 30000	0.88	1.11	0.23	26.02
			> 30000	1.54	2.11	0.58	37.63

Note: Increase = After – Before

Rear-end Crashes – Table Two
Statistical Modeling

Surrounding Land Use: Rural or Urban   Location Type: Business
Shoulder: Other

Intersection Characteristics				Number of Rear-end Crashes			
Speed	Lanes	Median	ADT (vpd)	Before	After	Increase	% Increase
≤ 45 mph	≤ 4	Divided	≤ 15000	0.45	0.96	0.50	111.13
			15000 to ≤ 30000	0.79	1.83	1.03	130.57
			> 30000	1.38	3.48	2.10	151.81
		Undivided	≤ 15000	0.53	0.77	0.24	44.53
			15000 to ≤ 30000	0.93	1.46	0.54	57.84
			> 30000	1.61	2.78	1.17	72.37
	> 4	Divided	≤ 15000	0.69	1.17	0.48	68.88
			15000 to ≤ 30000	1.21	2.22	1.02	84.43
			> 30000	2.10	4.24	2.13	101.42
		Undivided	≤ 15000	0.81	0.93	0.13	15.60
			15000 to ≤ 30000	1.41	1.78	0.37	26.25
			> 30000	2.46	3.39	0.93	37.88
> 45 mph	≤ 4	Divided	≤ 15000	0.28	0.83	0.54	192.18
			15000 to ≤ 30000	0.49	1.58	1.08	219.09
			> 30000	0.86	3.01	2.14	248.48
		Undivided	≤ 15000	0.33	0.66	0.33	100.01
			15000 to ≤ 30000	0.58	1.26	0.68	118.43
			> 30000	1.01	2.41	1.40	138.55
	> 4	Divided	≤ 15000	0.43	1.01	0.58	133.71
			15000 to ≤ 30000	0.75	1.92	1.17	155.23
			> 30000	1.31	3.66	2.35	178.74
		Undivided	≤ 15000	0.50	0.81	0.30	59.98
			15000 to ≤ 30000	0.88	1.54	0.66	74.72
			> 30000	1.54	2.93	1.39	90.81

Note: Increase = After – Before

Rear-end Crashes – Table Three
Statistical Modeling

Surrounding Land Use: Rural or Urban Location Type: Other
Shoulder: Paved

Intersection Characteristics				Number of Rear-end Crashes			
Speed	Lanes	Median	ADT (vpd)	Before	After	Increase	% Increase
≤ 45 mph	≤ 4	Divided	≤ 15000	0.35	0.69	0.34	94.90
			15000 to ≤ 30000	0.62	1.32	0.70	112.85
			> 30000	1.08	2.51	1.43	132.45
		Undivided	≤ 15000	0.41	0.55	0.14	33.42
			15000 to ≤ 30000	0.72	1.05	0.33	45.70
			> 30000	1.26	2.01	0.75	59.12
	> 4	Divided	≤ 15000	0.54	0.84	0.30	55.89
			15000 to ≤ 30000	0.94	1.60	0.66	70.25
			> 30000	1.64	3.06	1.41	85.93
		Undivided	≤ 15000	0.63	0.67	0.04	6.72
			15000 to ≤ 30000	1.10	1.28	0.18	16.54
			> 30000	1.92	2.45	0.52	27.28
> 45 mph	≤ 4	Divided	≤ 15000	0.22	0.60	0.38	169.72
			15000 to ≤ 30000	0.39	1.14	0.75	194.56
			> 30000	0.67	2.17	1.49	221.68
		Undivided	≤ 15000	0.26	0.48	0.22	84.63
			15000 to ≤ 30000	0.45	0.91	0.46	101.64
			> 30000	0.79	1.74	0.95	120.21
	> 4	Divided	≤ 15000	0.34	0.73	0.39	115.74
			15000 to ≤ 30000	0.59	1.39	0.80	135.61
			> 30000	1.03	2.64	1.61	157.31
		Undivided	≤ 15000	0.39	0.58	0.19	47.68
			15000 to ≤ 30000	0.69	1.11	0.42	61.28
			> 30000	1.20	2.11	0.91	76.14

Note: Increase = After – Before

Rear-end Crashes – Table Four
Statistical Modeling

Surrounding Land Use: Rural or Urban Location Type: Other
Shoulder: Other

Intersection Characteristics				Number of Rear-end Crashes			
Speed	Lanes	Median	ADT (vpd)	Before	After	Increase	% Increase
≤ 45 mph	≤ 4	Divided	≤ 15000	0.35	0.96	0.60	170.20
			15000 to ≤ 30000	0.62	1.83	1.21	195.09
			> 30000	1.08	3.48	2.40	222.26
		Undivided	≤ 15000	0.41	0.77	0.35	84.97
			15000 to ≤ 30000	0.72	1.46	0.74	102.00
			> 30000	1.26	2.78	1.52	120.60
	> 4	Divided	≤ 15000	0.54	1.17	0.63	116.13
			15000 to ≤ 30000	0.94	2.22	1.28	136.03
			> 30000	1.64	4.24	2.59	157.77
		Undivided	≤ 15000	0.63	0.93	0.30	47.95
			15000 to ≤ 30000	1.10	1.78	0.68	61.58
			> 30000	1.92	3.39	1.47	76.46
> 45 mph	≤ 4	Divided	≤ 15000	0.22	0.83	0.61	273.93
			15000 to ≤ 30000	0.39	1.58	1.19	308.37
			> 30000	0.67	3.01	2.33	345.98
		Undivided	≤ 15000	0.26	0.66	0.40	155.97
			15000 to ≤ 30000	0.45	1.26	0.81	179.55
			> 30000	0.79	2.41	1.62	205.29
	> 4	Divided	≤ 15000	0.34	1.01	0.67	199.10
			15000 to ≤ 30000	0.59	1.92	1.33	226.64
			> 30000	1.03	3.66	2.63	256.73
		Undivided	≤ 15000	0.39	0.81	0.41	104.75
			15000 to ≤ 30000	0.69	1.54	0.85	123.60
			> 30000	1.20	2.93	1.73	144.20

Note: Increase = After – Before

Other Crashes – Table One
Statistical Modeling

Surrounding Land Use: Urban Location Type: Business or Other
Shoulder: Paved or Other

Intersection Characteristics				Number of Other Crashes			
Speed	Lanes	Median	ADT (vpd)	Before	After	Increase	% Increase
≤ 45 mph	≤ 4	Divided	≤ 15000	0.83	0.55	-0.27	-33.17
			15000 to ≤ 30000	1.16	0.93	-0.23	-19.66
			> 30000	1.62	1.56	-0.06	-3.42
		Undivided	≤ 15000	0.63	0.55	-0.08	-12.22
			15000 to ≤ 30000	0.88	0.93	0.05	5.53
			> 30000	1.23	1.56	0.33	26.86
	> 4	Divided	≤ 15000	0.83	0.70	-0.13	-15.13
			15000 to ≤ 30000	1.16	1.18	0.02	2.03
			> 30000	1.62	1.98	0.37	22.65
		Undivided	≤ 15000	0.63	0.70	0.07	11.48
			15000 to ≤ 30000	0.88	1.18	0.30	34.02
			> 30000	1.23	1.98	0.75	61.11
> 45 mph	≤ 4	Divided	≤ 15000	0.54	0.55	0.02	3.38
			15000 to ≤ 30000	0.75	0.93	0.18	24.27
			> 30000	1.05	1.56	0.52	49.39
		Undivided	≤ 15000	0.41	0.55	0.15	35.78
			15000 to ≤ 30000	0.57	0.93	0.36	63.23
			> 30000	0.80	1.56	0.77	96.23
	> 4	Divided	≤ 15000	0.54	0.70	0.17	31.28
			15000 to ≤ 30000	0.75	1.18	0.43	57.82
			> 30000	1.05	1.98	0.94	89.72
		Undivided	≤ 15000	0.41	0.70	0.30	72.44
			15000 to ≤ 30000	0.57	1.18	0.61	107.30
			> 30000	0.80	1.98	1.19	149.20

Note: Increase = After – Before

Other Crashes – Table Two
Statistical Modeling

Surrounding Land Use: Rural Location Type: Business or Other
Shoulder: Paved or Other

Speed	Lanes	Median	ADT (vpd)	Before	After	Increase	% Increase
≤ 45 mph	≤ 4	Divided	≤ 15000	0.59	0.41	-0.18	-30.43
≤ 45 mph	≤ 4	Divided	15000 to ≤ 30000	0.82	0.69	-0.13	-16.36
≤ 45 mph	≤ 4	Divided	> 30000	1.15	1.16	0.01	0.54
≤ 45 mph	≤ 4	Undivided	≤ 15000	0.45	0.41	-0.04	-8.62
≤ 45 mph	≤ 4	Undivided	15000 to ≤ 30000	0.63	0.69	0.06	9.86
≤ 45 mph	≤ 4	Undivided	> 30000	0.88	1.16	0.28	32.06
≤ 45 mph	> 4	Divided	≤ 15000	0.59	0.52	-0.07	-11.64
≤ 45 mph	> 4	Divided	15000 to ≤ 30000	0.82	0.87	0.05	6.22
≤ 45 mph	> 4	Divided	> 30000	1.15	1.47	0.32	27.69
≤ 45 mph	> 4	Undivided	≤ 15000	0.45	0.52	0.07	16.06
≤ 45 mph	> 4	Undivided	15000 to ≤ 30000	0.63	0.87	0.25	39.51
≤ 45 mph	> 4	Undivided	> 30000	0.88	1.47	0.59	67.72
> 45 mph	≤ 4	Divided	≤ 15000	0.38	0.41	0.03	7.62
> 45 mph	≤ 4	Divided	15000 to ≤ 30000	0.53	0.69	0.16	29.37
> 45 mph	≤ 4	Divided	> 30000	0.74	1.16	0.41	55.52
> 45 mph	≤ 4	Undivided	≤ 15000	0.29	0.41	0.12	41.35
> 45 mph	≤ 4	Undivided	15000 to ≤ 30000	0.40	0.69	0.28	69.93
> 45 mph	≤ 4	Undivided	> 30000	0.57	1.16	0.59	104.28
> 45 mph	> 4	Divided	≤ 15000	0.38	0.52	0.14	36.67
> 45 mph	> 4	Divided	15000 to ≤ 30000	0.53	0.87	0.34	64.30
> 45 mph	> 4	Divided	> 30000	0.74	1.47	0.73	97.51
> 45 mph	> 4	Undivided	≤ 15000	0.29	0.52	0.23	79.52
> 45 mph	> 4	Undivided	15000 to ≤ 30000	0.40	0.87	0.47	115.80
> 45 mph	> 4	Undivided	> 30000	0.57	1.47	0.90	159.43

Note: Increase = After – Before

APPENDIX D

OPERATIONAL RESEARCH MODELING RESULTS

Total Crashes – Table One
Operational Research Modeling

Surrounding Land Use: Urban   Location Type: Business   Shoulder: Paved

Intersection Characteristics				Total Number of Crashes			
Speed	Lanes	Median	ADT (vpd)	Before	After	Increase	% Increase
≤ 45 mph	≤ 4	Divided	≤ 15000	4.07	3.78	-0.29	-7.04
			15000 to ≤ 30000	4.59	4.90	0.32	6.95
			> 30000	5.17	6.36	1.19	23.04
		Undivided	≤ 15000	2.73	2.47	-0.26	-9.65
			15000 to ≤ 30000	3.08	3.20	0.12	3.95
			> 30000	3.47	4.15	0.68	19.60
	> 4	Divided	≤ 15000	5.55	4.64	-0.91	-16.37
			15000 to ≤ 30000	6.25	6.02	-0.24	-3.78
			> 30000	7.05	7.80	0.75	10.70
		Undivided	≤ 15000	3.72	3.02	-0.70	-18.71
			15000 to ≤ 30000	4.19	3.92	-0.27	-6.48
			> 30000	4.73	5.08	0.36	7.60
> 45 mph	≤ 4	Divided	≤ 15000	3.39	3.78	0.40	11.71
			15000 to ≤ 30000	3.82	4.90	1.09	28.52
			> 30000	4.30	6.36	2.06	47.87
		Undivided	≤ 15000	2.27	2.47	0.19	8.58
			15000 to ≤ 30000	2.56	3.20	0.64	24.93
			> 30000	2.88	4.15	1.26	43.73
	> 4	Divided	≤ 15000	4.62	4.64	0.02	0.50
			15000 to ≤ 30000	5.20	6.02	0.81	15.63
			> 30000	5.86	7.80	1.94	33.03
		Undivided	≤ 15000	3.10	3.02	-0.07	-2.31
			15000 to ≤ 30000	3.49	3.92	0.43	12.39
			> 30000	3.93	5.08	1.15	29.31

Note: Increase = After – Before

Total Crashes – Table Two
Operational Research Modeling

Surrounding Land Use: Urban　　Location Type: Business　　Shoulder: Other

Intersection Characteristics				Total Number of Crashes			
Speed	Lanes	Median	ADT (vpd)	Before	After	Increase	% Increase
≤ 45 mph	≤ 4	Divided	≤ 15000	4.50	4.22	-0.28	-6.16
			15000 to ≤ 30000	5.07	5.47	0.40	7.96
			> 30000	5.71	7.09	1.38	24.21
		Undivided	≤ 15000	3.01	2.75	-0.26	-8.79
			15000 to ≤ 30000	3.40	3.57	0.17	4.94
			> 30000	3.83	4.62	0.79	20.74
	> 4	Divided	≤ 15000	6.13	5.17	-0.95	-15.58
			15000 to ≤ 30000	6.91	6.71	-0.20	-2.87
			> 30000	7.78	8.70	0.91	11.75
		Undivided	≤ 15000	4.11	3.37	-0.74	-17.94
			15000 to ≤ 30000	4.63	4.37	-0.26	-5.59
			> 30000	5.22	5.67	0.45	8.62
> 45 mph	≤ 4	Divided	≤ 15000	3.74	4.22	0.48	12.77
			15000 to ≤ 30000	4.22	5.47	1.25	29.75
			> 30000	4.75	7.09	2.34	49.27
		Undivided	≤ 15000	2.51	2.75	0.24	9.62
			15000 to ≤ 30000	2.83	3.57	0.74	26.12
			> 30000	3.19	4.62	1.44	45.10
	> 4	Divided	≤ 15000	5.10	5.17	0.07	1.46
			15000 to ≤ 30000	5.75	6.71	0.96	16.73
			> 30000	6.48	8.70	2.22	34.30
		Undivided	≤ 15000	3.42	3.37	-0.05	-1.38
			15000 to ≤ 30000	3.85	4.37	0.52	13.46
			> 30000	4.34	5.67	1.33	30.54

Note: Increase = After – Before

Total Crashes – Table Three
Operational Research Modeling

Surrounding Land Use: Urban   Location Type: Other   Shoulder: Paved

Intersection Characteristics				Total Number of Crashes			
Speed	Lanes	Median	ADT (vpd)	Before	After	Increase	% Increase
≤ 45 mph	≤ 4	Divided	≤ 15000	3.92	3.15	-0.77	-19.70
			15000 to ≤ 30000	4.42	4.08	-0.34	-7.62
			> 30000	4.98	5.29	0.31	6.29
		Undivided	≤ 15000	2.63	2.05	-0.58	-21.95
			15000 to ≤ 30000	2.96	2.66	-0.30	-10.20
			> 30000	3.34	3.45	0.11	3.31
	> 4	Divided	≤ 15000	5.34	3.86	-1.48	-27.76
			15000 to ≤ 30000	6.02	5.00	-1.02	-16.89
			> 30000	6.79	6.49	-0.30	-4.38
		Undivided	≤ 15000	3.58	2.52	-1.07	-29.78
			15000 to ≤ 30000	4.04	3.26	-0.78	-19.21
			> 30000	4.55	4.23	-0.32	-7.06
> 45 mph	≤ 4	Divided	≤ 15000	3.26	3.15	-0.11	-3.50
			15000 to ≤ 30000	3.68	4.08	0.41	11.02
			> 30000	4.14	5.29	1.15	27.73
		Undivided	≤ 15000	2.19	2.05	-0.14	-6.20
			15000 to ≤ 30000	2.46	2.66	0.20	7.91
			> 30000	2.78	3.45	0.67	24.15
	> 4	Divided	≤ 15000	4.45	3.86	-0.59	-13.18
			15000 to ≤ 30000	5.01	5.00	-0.01	-0.12
			> 30000	5.65	6.49	0.84	14.91
		Undivided	≤ 15000	2.98	2.52	-0.47	-15.62
			15000 to ≤ 30000	3.36	3.26	-0.10	-2.91
			> 30000	3.79	4.23	0.44	11.70

Note: Increase = After – Before

Total Crashes – Table Four
Operational Research Modeling

Surrounding Land Use: Urban   Location Type: Other   Shoulder: Other

Intersection Characteristics				Total Number of Crashes			
Speed	Lanes	Median	ADT (vpd)	Before	After	Increase	% Increase
≤ 45 mph	≤ 4	Divided	≤ 15000	4.33	3.51	-0.82	-18.94
			15000 to ≤ 30000	4.88	4.55	-0.33	-6.74
			> 30000	5.50	5.90	0.40	7.30
		Undivided	≤ 15000	2.90	2.29	-0.62	-21.21
			15000 to ≤ 30000	3.27	2.97	-0.31	-9.35
			> 30000	3.69	3.85	0.16	4.29
	> 4	Divided	≤ 15000	5.90	4.30	-1.60	-27.07
			15000 to ≤ 30000	6.65	5.58	-1.07	-16.10
			> 30000	7.50	7.24	-0.26	-3.47
		Undivided	≤ 15000	3.96	2.81	-1.15	-29.11
			15000 to ≤ 30000	4.46	3.64	-0.82	-18.45
			> 30000	5.03	4.72	-0.31	-6.17
> 45 mph	≤ 4	Divided	≤ 15000	3.60	3.51	-0.09	-2.58
			15000 to ≤ 30000	4.06	4.55	0.49	12.08
			> 30000	4.58	5.90	1.32	28.95
		Undivided	≤ 15000	2.42	2.29	-0.13	-5.31
			15000 to ≤ 30000	2.72	2.97	0.24	8.94
			> 30000	3.07	3.85	0.78	25.34
	> 4	Divided	≤ 15000	4.91	4.30	-0.61	-12.36
			15000 to ≤ 30000	5.53	5.58	0.05	0.83
			> 30000	6.24	7.24	1.00	16.01
		Undivided	≤ 15000	3.29	2.81	-0.49	-14.81
			15000 to ≤ 30000	3.71	3.64	-0.07	-1.99
			> 30000	4.18	4.72	0.53	12.76

Note: Increase = After – Before

Total Crashes – Table Five
Operational Research Modeling

Surrounding Land Use: Rural Location Type: Business Shoulder: Paved

Intersection Characteristics				Total Number of Crashes			
Speed	Lanes	Median	ADT (vpd)	Before	After	Increase	% Increase
≤ 45 mph	≤ 4	Divided	≤ 15000	3.36	3.26	-0.10	-3.06
			15000 to ≤ 30000	3.79	4.23	0.44	11.53
			> 30000	4.27	5.48	1.21	28.31
		Undivided	≤ 15000	2.26	2.13	-0.13	-5.78
			15000 to ≤ 30000	2.54	2.76	0.21	8.41
			> 30000	2.87	3.57	0.71	24.72
	> 4	Divided	≤ 15000	4.59	4.00	-0.59	-12.79
			15000 to ≤ 30000	5.17	5.19	0.02	0.34
			> 30000	5.83	6.73	0.90	15.44
		Undivided	≤ 15000	3.08	2.61	-0.47	-15.23
			15000 to ≤ 30000	3.47	3.38	-0.09	-2.47
			> 30000	3.91	4.38	0.48	12.21
> 45 mph	≤ 4	Divided	≤ 15000	2.80	3.26	0.46	16.50
			15000 to ≤ 30000	3.16	4.23	1.07	34.03
			> 30000	3.56	5.48	1.93	54.20
		Undivided	≤ 15000	1.88	2.13	0.25	13.24
			15000 to ≤ 30000	2.12	2.76	0.64	30.28
			> 30000	2.38	3.57	1.19	49.89
	> 4	Divided	≤ 15000	3.82	4.00	0.18	4.81
			15000 to ≤ 30000	4.30	5.19	0.89	20.58
			> 30000	4.85	6.73	1.88	38.73
		Undivided	≤ 15000	2.56	2.61	0.05	1.87
			15000 to ≤ 30000	2.88	3.38	0.50	17.21
			> 30000	3.25	4.38	1.13	34.85

Note: Increase = After – Before

Total Crashes – Table Six
Operational Research Modeling

Surrounding Land Use: Rural Location Type: Business Shoulder: Other

Intersection Characteristics				Total Number of Crashes			
Speed	Lanes	Median	ADT (vpd)	Before	After	Increase	% Increase
≤ 45 mph	≤ 4	Divided	≤ 15000	3.72	3.64	-0.08	-2.14
			15000 to ≤ 30000	4.19	4.72	0.53	12.59
			> 30000	4.72	6.11	1.39	29.53
		Undivided	≤ 15000	2.49	2.37	-0.12	-4.88
			15000 to ≤ 30000	2.81	3.07	0.27	9.44
			> 30000	3.17	3.99	0.82	25.91
	> 4	Divided	≤ 15000	5.07	4.46	-0.61	-11.96
			15000 to ≤ 30000	5.71	5.78	0.07	1.29
			> 30000	6.44	7.50	1.06	16.54
		Undivided	≤ 15000	3.40	2.91	-0.49	-14.42
			15000 to ≤ 30000	3.83	3.77	-0.06	-1.54
			> 30000	4.32	4.89	0.57	13.27
> 45 mph	≤ 4	Divided	≤ 15000	3.09	3.64	0.54	17.61
			15000 to ≤ 30000	3.48	4.72	1.23	35.31
			> 30000	3.93	6.11	2.19	55.67
		Undivided	≤ 15000	2.07	2.37	0.30	14.31
			15000 to ≤ 30000	2.34	3.07	0.74	31.52
			> 30000	2.63	3.99	1.35	51.31
	> 4	Divided	≤ 15000	4.22	4.46	0.24	5.80
			15000 to ≤ 30000	4.75	5.78	1.03	21.73
			> 30000	5.35	7.50	2.14	40.05
		Undivided	≤ 15000	2.83	2.91	0.08	2.84
			15000 to ≤ 30000	3.19	3.77	0.58	18.32
			> 30000	3.59	4.89	1.30	36.13

Note: Increase = After – Before

Total Crashes – Table Seven
Operational Research Modeling

Surrounding Land Use: Rural   Location Type: Other   Shoulder: Paved

Intersection Characteristics				Total Number of Crashes			
Speed	Lanes	Median	ADT (vpd)	Before	After	Increase	% Increase
≤ 45 mph	≤ 4	Divided	≤ 15000	3.24	2.71	-0.53	-16.26
			15000 to ≤ 30000	3.65	3.52	-0.13	-3.66
			> 30000	4.11	4.56	0.45	10.84
		Undivided	≤ 15000	2.17	1.77	-0.40	-18.61
			15000 to ≤ 30000	2.45	2.29	-0.16	-6.36
			> 30000	2.76	2.97	0.21	7.74
	> 4	Divided	≤ 15000	4.42	3.33	-1.09	-24.67
			15000 to ≤ 30000	4.98	4.32	-0.66	-13.33
			> 30000	5.61	5.59	-0.02	-0.28
		Undivided	≤ 15000	2.96	2.17	-0.79	-26.77
			15000 to ≤ 30000	3.34	2.81	-0.53	-15.75
			> 30000	3.76	3.64	-0.12	-3.07
> 45 mph	≤ 4	Divided	≤ 15000	2.70	2.71	0.02	0.63
			15000 to ≤ 30000	3.04	3.52	0.48	15.78
			> 30000	3.42	4.56	1.14	33.20
		Undivided	≤ 15000	1.81	1.77	-0.04	-2.19
			15000 to ≤ 30000	2.04	2.29	0.26	12.54
			> 30000	2.30	2.97	0.68	29.47
	> 4	Divided	≤ 15000	3.68	3.33	-0.35	-9.47
			15000 to ≤ 30000	4.14	4.31	0.17	4.16
			> 30000	4.67	5.60	0.93	19.84
		Undivided	≤ 15000	2.47	2.17	-0.30	-12.00
			15000 to ≤ 30000	2.78	2.81	0.03	1.24
			> 30000	3.13	3.65	0.52	16.48

Note: Increase = After – Before

Total Crashes – Table Eight
Operational Research Modeling

Surrounding Land Use: Rural　　Location Type: Other　　Shoulder: Other

Intersection Characteristics				Total Number of Crashes			
Speed	Lanes	Median	ADT (vpd)	Before	After	Increase	% Increase
≤ 45 mph	≤ 4	Divided	≤ 15000	3.58	3.03	-0.55	-15.47
			15000 to ≤ 30000	4.03	3.92	-0.11	-2.74
			> 30000	4.55	5.09	0.54	11.89
		Undivided	≤ 15000	2.40	1.97	-0.43	-17.83
			15000 to ≤ 30000	2.70	2.56	-0.15	-5.47
			> 30000	3.05	3.32	0.27	8.76
	> 4	Divided	≤ 15000	4.88	3.71	-1.17	-23.95
			15000 to ≤ 30000	5.50	4.81	-0.69	-12.50
			> 30000	6.20	6.24	0.04	0.67
		Undivided	≤ 15000	3.27	2.42	-0.85	-26.08
			15000 to ≤ 30000	3.69	3.14	-0.55	-14.95
			> 30000	4.16	4.07	-0.09	-2.15
> 45 mph	≤ 4	Divided	≤ 15000	2.98	3.03	0.05	1.59
			15000 to ≤ 30000	3.36	3.92	0.57	16.88
			> 30000	3.78	5.09	1.30	34.47
		Undivided	≤ 15000	2.00	1.97	-0.03	-1.25
			15000 to ≤ 30000	2.25	2.56	0.31	13.61
			> 30000	2.54	3.32	0.78	30.70
	> 4	Divided	≤ 15000	4.06	3.71	-0.35	-8.60
			15000 to ≤ 30000	4.58	4.81	0.24	5.15
			> 30000	5.16	6.24	1.08	20.98
		Undivided	≤ 15000	2.72	2.42	-0.30	-11.16
			15000 to ≤ 30000	3.07	3.14	0.07	2.21
			> 30000	3.46	4.07	0.61	17.59

Note: Increase = After – Before

Angle Crashes – Table One
Operational Research Modeling

Surrounding Land Use: Urban   Location Type: Business   Shoulder: Paved
Speed: ≤ 45 mph or > 45 mph

Intersection Characteristics			Number of Angle Crashes			
Lanes	Median	ADT (vpd)	Before	After	Increase	% Increase
≤ 4	Divided	≤ 15000	1.06	0.72	-0.33	-31.53
		15000 to ≤ 30000	1.08	0.90	-0.18	-16.55
		> 30000	1.11	1.13	0.02	1.70
	Undivided	≤ 15000	0.63	0.59	-0.03	-5.53
		15000 to ≤ 30000	0.64	0.74	0.10	15.13
		> 30000	0.66	0.92	0.26	40.30
> 4	Divided	≤ 15000	1.06	0.92	-0.14	-13.09
		15000 to ≤ 30000	1.08	1.15	0.06	5.92
		> 30000	1.11	1.43	0.32	29.08
	Undivided	≤ 15000	0.63	0.75	0.12	19.90
		15000 to ≤ 30000	0.64	0.94	0.30	46.13
		> 30000	0.66	1.17	0.51	78.08

Note: Increase = After – Before

Angle Crashes – Table Two
Operational Research Modeling

Surrounding Land Use: Urban Location Type: Business Shoulder: Other
Speed: ≤ 45 mph or > 45 mph

Intersection Characteristics			Number of Angle Crashes			
Lanes	Median	ADT (vpd)	Before	After	Increase	% Increase
≤ 4	Divided	≤ 15000	1.55	0.72	-0.82	-53.13
		15000 to ≤ 30000	1.58	0.90	-0.68	-42.88
		> 30000	1.62	1.13	-0.49	-30.39
	Undivided	≤ 15000	0.92	0.59	-0.32	-35.34
		15000 to ≤ 30000	0.94	0.74	-0.20	-21.20
		> 30000	0.96	0.92	-0.04	-3.97
> 4	Divided	≤ 15000	1.55	0.92	-0.63	-40.51
		15000 to ≤ 30000	1.58	1.15	-0.44	-27.50
		> 30000	1.62	1.43	-0.19	-11.65
	Undivided	≤ 15000	0.92	0.75	-0.16	-17.93
		15000 to ≤ 30000	0.94	0.94	0.00	0.02
		> 30000	0.96	1.17	0.21	21.89

Note: Increase = After – Before

Angle Crashes – Table Three
Operational Research Modeling

Surrounding Land Use: Urban Location Type: Other Shoulder: Paved
Speed: ≤ 45 mph or > 45 mph

Intersection Characteristics			Number of Angle Crashes			
Lanes	Median	ADT (vpd)	Before	After	Increase	% Increase
≤ 4	Divided	≤ 15000	1.06	0.57	-0.48	-45.81
		15000 to ≤ 30000	1.08	0.72	-0.37	-33.96
		> 30000	1.11	0.89	-0.22	-19.52
	Undivided	≤ 15000	0.63	0.47	-0.16	-25.24
		15000 to ≤ 30000	0.64	0.59	-0.06	-8.90
		> 30000	0.66	0.73	0.07	11.03
> 4	Divided	≤ 15000	1.06	0.73	-0.33	-31.22
		15000 to ≤ 30000	1.08	0.91	-0.18	-16.18
		> 30000	1.11	1.13	0.02	2.15
	Undivided	≤ 15000	0.63	0.60	-0.03	-5.12
		15000 to ≤ 30000	0.64	0.74	0.10	15.63
		> 30000	0.66	0.93	0.27	40.92

Note: Increase = After – Before

Angle Crashes – Table Four
Operational Research Modeling

Surrounding Land Use: Urban   Location Type: Other   Shoulder: Other
Speed: ≤ 45 mph or > 45 mph

Intersection Characteristics			Number of Angle Crashes			
Lanes	Median	ADT (vpd)	Before	After	Increase	% Increase
≤ 4	Divided	≤ 15000	1.55	0.57	-0.97	-62.91
		15000 to ≤ 30000	1.58	0.72	-0.87	-54.80
		> 30000	1.62	0.89	-0.73	-44.92
	Undivided	≤ 15000	0.92	0.47	-0.45	-48.83
		15000 to ≤ 30000	0.94	0.59	-0.35	-37.64
		> 30000	0.96	0.73	-0.23	-24.00
> 4	Divided	≤ 15000	1.55	0.73	-0.82	-52.93
		15000 to ≤ 30000	1.58	0.91	-0.67	-42.63
		> 30000	1.62	1.13	-0.49	-30.08
	Undivided	≤ 15000	0.92	0.60	-0.32	-35.05
		15000 to ≤ 30000	0.94	0.74	-0.20	-20.85
		> 30000	0.96	0.93	-0.03	-3.54

Note: Increase = After – Before

Angle Crashes – Table Five
Operational Research Modeling

Surrounding Land Use: Rural   Location Type: Business   Shoulder: Paved
Speed: ≤ 45 mph or > 45 mph

Intersection Characteristics			Number of Angle Crashes			
Lanes	Median	ADT (vpd)	Before	After	Increase	% Increase
≤ 4	Divided	≤ 15000	0.85	0.72	-0.13	-14.87
		15000 to ≤ 30000	0.87	0.90	0.03	3.74
		> 30000	0.89	1.13	0.24	26.43
	Undivided	≤ 15000	0.50	0.59	0.09	17.44
		15000 to ≤ 30000	0.52	0.74	0.22	43.13
		> 30000	0.53	0.92	0.39	74.43
> 4	Divided	≤ 15000	0.85	0.92	0.07	8.05
		15000 to ≤ 30000	0.87	1.15	0.28	31.68
		> 30000	0.89	1.43	0.54	60.47
	Undivided	≤ 15000	0.50	0.75	0.25	49.06
		15000 to ≤ 30000	0.52	0.94	0.42	81.66
		> 30000	0.53	1.17	0.64	121.39

Note: Increase = After – Before

Angle Crashes – Table Six
Operational Research Modeling

Surrounding Land Use: Rural   Location Type: Business   Shoulder: Other
Speed: ≤ 45 mph or > 45 mph

Intersection Characteristics			Number of Angle Crashes			
Lanes	Median	ADT (vpd)	Before	After	Increase	% Increase
≤ 4	Divided	≤ 15000	1.24	0.72	-0.52	-41.73
		15000 to ≤ 30000	1.27	0.90	-0.37	-28.99
		> 30000	1.30	1.13	-0.18	-13.46
	Undivided	≤ 15000	0.74	0.59	-0.14	-19.61
		15000 to ≤ 30000	0.75	0.74	-0.02	-2.03
		> 30000	0.77	0.92	0.15	19.39
> 4	Divided	≤ 15000	1.24	0.92	-0.32	-26.04
		15000 to ≤ 30000	1.27	1.15	-0.13	-9.87
		> 30000	1.30	1.43	0.13	9.84
	Undivided	≤ 15000	0.74	0.75	0.01	2.03
		15000 to ≤ 30000	0.75	0.94	0.18	24.34
		> 30000	0.77	1.17	0.40	51.53

Note: Increase = After – Before

Angle Crashes – Table Seven
Operational Research Modeling

Surrounding Land Use: Rural Location Type: Other Shoulder: Paved
Speed: ≤ 45 mph or > 45 mph

Intersection Characteristics			Number of Angle Crashes			
Lanes	Median	ADT (vpd)	Before	After	Increase	% Increase
≤ 4	Divided	≤ 15000	0.85	0.57	-0.28	-32.64
		15000 to ≤ 30000	0.87	0.72	-0.16	-17.90
		> 30000	0.89	0.89	0.00	0.05
	Undivided	≤ 15000	0.50	0.47	-0.04	-7.06
		15000 to ≤ 30000	0.52	0.59	0.07	13.26
		> 30000	0.53	0.73	0.20	38.03
> 4	Divided	≤ 15000	0.85	0.73	-0.12	-14.50
		15000 to ≤ 30000	0.87	0.91	0.04	4.20
		> 30000	0.89	1.13	0.24	26.99
	Undivided	≤ 15000	0.50	0.60	0.09	17.96
		15000 to ≤ 30000	0.52	0.74	0.23	43.76
		> 30000	0.53	0.93	0.40	75.20

Note: Increase = After – Before

Angle Crashes – Table Eight
Operational Research Modeling

Surrounding Land Use: Rural Location Type: Other Shoulder: Other
Speed: ≤ 45 mph or > 45 mph

Intersection Characteristics			Number of Angle Crashes			
Lanes	Median	ADT (vpd)	Before	After	Increase	% Increase
≤ 4	Divided	≤ 15000	1.24	0.57	-0.67	-53.89
		15000 to ≤ 30000	1.27	0.72	-0.56	-43.81
		> 30000	1.30	0.89	-0.41	-31.52
	Undivided	≤ 15000	0.74	0.47	-0.27	-36.39
		15000 to ≤ 30000	0.75	0.59	-0.17	-22.48
		> 30000	0.77	0.73	-0.04	-5.52
> 4	Divided	≤ 15000	1.24	0.73	-0.52	-41.48
		15000 to ≤ 30000	1.27	0.91	-0.36	-28.68
		> 30000	1.30	1.13	-0.17	-13.08
	Undivided	≤ 15000	0.74	0.60	-0.14	-19.26
		15000 to ≤ 30000	0.75	0.74	-0.01	-1.60
		> 30000	0.77	0.93	0.15	19.92

Note: Increase = After – Before

Left-turn Crashes – Table One
Operational Research Modeling

Surrounding Land Use: Urban or Rural — Location Type: Business
Shoulder: Paved — Speed: ≤ 45 mph or > 45 mph

Intersection Characteristics			Number of Left-turn Crashes			
Lanes	Median	ADT (vpd)	Before	After	Increase	% Increase
≤ 4	Divided	≤ 15000	0.82	0.53	-0.29	-35.46
		15000 to ≤ 30000	0.98	0.70	-0.29	-29.18
		> 30000	1.18	0.91	-0.26	-22.29
	Undivided	≤ 15000	0.55	0.28	-0.27	-49.68
		15000 to ≤ 30000	0.66	0.36	-0.29	-44.78
		> 30000	0.78	0.47	-0.31	-39.40
> 4	Divided	≤ 15000	1.17	0.73	-0.44	-37.92
		15000 to ≤ 30000	1.40	0.95	-0.45	-31.88
		> 30000	1.67	1.25	-0.42	-25.25
	Undivided	≤ 15000	0.78	0.38	-0.40	-51.59
		15000 to ≤ 30000	0.93	0.49	-0.44	-46.88
		> 30000	1.11	0.65	-0.46	-41.71

Note: Increase = After – Before

Left-turn Crashes – Table Two
Operational Research Modeling

Surrounding Land Use: Urban or Rural — Location Type: Business
Shoulder: Other — Speed: ≤ 45 mph or > 45 mph

<table>
<tr><th colspan="3">Intersection Characteristics</th><th colspan="4">Number of Left-turn Crashes</th></tr>
<tr><th>Lanes</th><th>Median</th><th>ADT (vpd)</th><th>Before</th><th>After</th><th>Increase</th><th>% Increase</th></tr>
<tr><td rowspan="6">≤ 4</td><td rowspan="3">Divided</td><td>≤ 15000</td><td>0.94</td><td>0.70</td><td>-0.24</td><td>-25.32</td></tr>
<tr><td>15000 to ≤ 30000</td><td>1.12</td><td>0.92</td><td>-0.20</td><td>-18.05</td></tr>
<tr><td>> 30000</td><td>1.34</td><td>1.20</td><td>-0.13</td><td>-10.08</td></tr>
<tr><td rowspan="3">Undivided</td><td>≤ 15000</td><td>0.62</td><td>0.36</td><td>-0.26</td><td>-41.77</td></tr>
<tr><td>15000 to ≤ 30000</td><td>0.75</td><td>0.48</td><td>-0.27</td><td>-36.10</td></tr>
<tr><td>> 30000</td><td>0.89</td><td>0.62</td><td>-0.27</td><td>-29.88</td></tr>
<tr><td rowspan="6">> 4</td><td rowspan="3">Divided</td><td>≤ 15000</td><td>1.33</td><td>0.96</td><td>-0.37</td><td>-28.17</td></tr>
<tr><td>15000 to ≤ 30000</td><td>1.59</td><td>1.25</td><td>-0.34</td><td>-21.18</td></tr>
<tr><td>> 30000</td><td>1.90</td><td>1.64</td><td>-0.26</td><td>-13.50</td></tr>
<tr><td rowspan="3">Undivided</td><td>≤ 15000</td><td>0.89</td><td>0.50</td><td>-0.39</td><td>-43.99</td></tr>
<tr><td>15000 to ≤ 30000</td><td>1.06</td><td>0.65</td><td>-0.41</td><td>-38.54</td></tr>
<tr><td>> 30000</td><td>1.27</td><td>0.85</td><td>-0.41</td><td>-32.55</td></tr>
</table>

Note: Increase = After – Before

Left-turn Crashes – Table Three
Operational Research Modeling

Surrounding Land Use: Urban or Rural
Shoulder: Paved

Location Type: Other
Speed: ≤ 45 mph or > 45 mph

Intersection Characteristics			Number of Left-turn Crashes			
Lanes	Median	ADT (vpd)	Before	After	Increase	% Increase
≤ 4	Divided	≤ 15000	0.68	0.49	-0.19	-28.40
		15000 to ≤ 30000	0.81	0.64	-0.17	-21.43
		> 30000	0.97	0.84	-0.13	-13.78
	Undivided	≤ 15000	0.45	0.25	-0.20	-44.17
		15000 to ≤ 30000	0.54	0.33	-0.21	-38.73
		> 30000	0.65	0.43	-0.21	-32.77
> 4	Divided	≤ 15000	0.97	0.66	-0.30	-31.13
		15000 to ≤ 30000	1.15	0.87	-0.28	-24.42
		> 30000	1.38	1.14	-0.24	-17.07
	Undivided	≤ 15000	0.64	0.35	-0.30	-46.29
		15000 to ≤ 30000	0.77	0.45	-0.32	-41.07
		> 30000	0.92	0.59	-0.32	-35.33

Note: Increase = After – Before

Left-turn Crashes – Table four
Operational Research Modeling

Surrounding Land Use: Urban or Rural — Location Type: Other
Shoulder: Other — Speed: ≤ 45 mph or > 45 mph

Intersection Characteristics			Number of Left-turn Crashes			
Lanes	Median	ADT (vpd)	Before	After	Increase	% Increase
≤ 4	Divided	≤ 15000	0.77	0.64	-0.13	-17.15
		15000 to ≤ 30000	0.92	0.84	-0.08	-9.08
		> 30000	1.10	1.10	0.00	-0.23
	Undivided	≤ 15000	0.51	0.33	-0.18	-35.39
		15000 to ≤ 30000	0.62	0.44	-0.18	-29.11
		> 30000	0.74	0.57	-0.16	-22.21
> 4	Divided	≤ 15000	1.10	0.87	-0.22	-20.30
		15000 to ≤ 30000	1.31	1.15	-0.16	-12.55
		> 30000	1.57	1.50	-0.06	-4.03
	Undivided	≤ 15000	0.73	0.45	-0.28	-37.85
		15000 to ≤ 30000	0.87	0.60	-0.28	-31.81
		> 30000	1.04	0.78	-0.26	-25.17

Note: Increase = After – Before

Rear-end Crashes – Table One
Operational Research Modeling

Surrounding Land Use: Urban or Rural | Location Type: Business
Shoulder: Paved

Intersection Characteristics				Number of Rear-end Crashes			
Speed	Lanes	Median	ADT (vpd)	Before	After	Increase	% Increase
≤ 45 mph	≤ 4	Divided	≤ 15000	0.54	0.91	0.37	68.00
			15000 to ≤ 30000	0.90	1.44	0.54	60.47
			> 30000	1.49	2.29	0.80	53.28
		Undivided	≤ 15000	0.57	0.55	-0.02	-3.42
			15000 to ≤ 30000	0.94	0.87	-0.07	-7.75
			> 30000	1.57	1.38	-0.19	-11.88
	> 4	Divided	≤ 15000	0.66	1.03	0.37	55.21
			15000 to ≤ 30000	1.10	1.63	0.53	48.25
			> 30000	1.83	2.59	0.76	41.60
		Undivided	≤ 15000	0.70	0.62	-0.07	-10.77
			15000 to ≤ 30000	1.15	0.98	-0.17	-14.77
			> 30000	1.92	1.56	-0.36	-18.59
> 45 mph	≤ 4	Divided	≤ 15000	0.39	0.97	0.59	151.65
			15000 to ≤ 30000	0.64	1.54	0.90	140.37
			> 30000	1.06	2.44	1.38	129.59
		Undivided	≤ 15000	0.41	0.59	0.18	44.67
			15000 to ≤ 30000	0.67	0.93	0.26	38.19
			> 30000	1.12	1.48	0.36	31.99
	> 4	Divided	≤ 15000	0.47	1.10	0.63	132.49
			15000 to ≤ 30000	0.78	1.74	0.96	122.06
			> 30000	1.30	2.76	1.46	112.11
		Undivided	≤ 15000	0.50	0.66	0.17	33.65
			15000 to ≤ 30000	0.82	1.05	0.23	27.66
			> 30000	1.37	1.67	0.30	21.94

Note: Increase = After – Before

Rear-end Crashes – Table Two
Operational Research Modeling

Surrounding Land Use: Urban or Rural — Location Type: Business
Shoulder: Other

Intersection Characteristics				Number of Rear-end Crashes			
Speed	Lanes	Median	ADT (vpd)	Before	After	Increase	% Increase
≤ 45 mph	≤ 4	Divided	≤ 15000	0.54	1.22	0.68	125.48
			15000 to ≤ 30000	0.90	1.94	1.04	115.37
			> 30000	1.49	3.07	1.58	105.72
		Undivided	≤ 15000	0.57	0.74	0.17	29.63
			15000 to ≤ 30000	0.94	1.17	0.22	23.82
			> 30000	1.57	1.85	0.29	18.27
	> 4	Divided	≤ 15000	0.66	1.38	0.72	108.31
			15000 to ≤ 30000	1.10	2.19	1.09	98.97
			> 30000	1.83	3.47	1.64	90.05
		Undivided	≤ 15000	0.70	0.83	0.14	19.75
			15000 to ≤ 30000	1.15	1.32	0.17	14.39
			> 30000	1.92	2.10	0.18	9.26
> 45 mph	≤ 4	Divided	≤ 15000	0.39	1.30	0.92	237.74
			15000 to ≤ 30000	0.64	2.07	1.43	222.60
			> 30000	1.06	3.28	2.22	208.14
		Undivided	≤ 15000	0.41	0.79	0.38	94.17
			15000 to ≤ 30000	0.67	1.25	0.58	85.46
			> 30000	1.12	1.98	0.86	77.15
	> 4	Divided	≤ 15000	0.47	1.47	1.00	212.02
			15000 to ≤ 30000	0.78	2.34	1.55	198.03
			> 30000	1.30	3.71	2.40	184.67
		Undivided	≤ 15000	0.50	0.89	0.39	79.38
			15000 to ≤ 30000	0.82	1.41	0.59	71.34
			> 30000	1.37	2.24	0.87	63.66

Note: Increase = After – Before

Rear-end Crashes – Table Three
Operational Research Modeling

Surrounding Land Use: Urban or Rural — Location Type: Other
Shoulder: Paved

Intersection Characteristics				Number of Rear-end Crashes			
Speed	Lanes	Median	ADT (vpd)	Before	After	Increase	% Increase
≤ 45 mph	≤ 4	Divided	≤ 15000	0.38	0.91	0.53	137.33
			15000 to ≤ 30000	0.64	1.44	0.81	126.69
			> 30000	1.06	2.29	1.23	116.53
		Undivided	≤ 15000	0.40	0.55	0.15	36.44
			15000 to ≤ 30000	0.67	0.87	0.20	30.32
			> 30000	1.11	1.38	0.27	24.48
	> 4	Divided	≤ 15000	0.47	1.03	0.56	119.26
			15000 to ≤ 30000	0.78	1.63	0.85	109.43
			> 30000	1.29	2.59	1.29	100.04
		Undivided	≤ 15000	0.49	0.62	0.13	26.05
			15000 to ≤ 30000	0.82	0.98	0.17	20.40
			> 30000	1.36	1.56	0.20	15.00
> 45 mph	≤ 4	Divided	≤ 15000	0.27	0.97	0.70	255.50
			15000 to ≤ 30000	0.45	1.54	1.09	239.56
			> 30000	0.75	2.44	1.69	224.34
		Undivided	≤ 15000	0.29	0.59	0.30	104.37
			15000 to ≤ 30000	0.48	0.93	0.45	95.21
			> 30000	0.79	1.48	0.68	86.46
	> 4	Divided	≤ 15000	0.33	1.10	0.76	228.42
			15000 to ≤ 30000	0.56	1.74	1.19	213.70
			> 30000	0.92	2.76	1.84	199.64
		Undivided	≤ 15000	0.35	0.66	0.31	88.81
			15000 to ≤ 30000	0.58	1.05	0.47	80.35
			> 30000	0.97	1.67	0.70	72.26

Note: Increase = After – Before

Rear-end Crashes – Table Four
Operational Research Modeling

Surrounding Land Use: Urban or Rural | Location Type: Other
Shoulder: Other

Intersection Characteristics				Number of Rear-end Crashes			
Speed	Lanes	Median	ADT (vpd)	Before	After	Increase	% Increase
≤ 45 mph	≤ 4	Divided	≤ 15000	0.38	1.22	0.84	218.53
			15000 to ≤ 30000	0.64	1.94	1.30	204.25
			> 30000	1.06	3.07	2.02	190.61
		Undivided	≤ 15000	0.40	0.74	0.33	83.12
			15000 to ≤ 30000	0.67	1.17	0.50	74.91
			> 30000	1.11	1.85	0.74	67.07
	> 4	Divided	≤ 15000	0.47	1.38	0.91	194.27
			15000 to ≤ 30000	0.78	2.19	1.41	181.08
			> 30000	1.29	3.47	2.18	168.48
		Undivided	≤ 15000	0.49	0.83	0.34	69.17
			15000 to ≤ 30000	0.82	1.32	0.50	61.59
			> 30000	1.36	2.10	0.74	54.35
> 45 mph	≤ 4	Divided	≤ 15000	0.27	1.30	1.03	377.12
			15000 to ≤ 30000	0.45	2.07	1.61	355.73
			> 30000	0.75	3.28	2.53	335.30
		Undivided	≤ 15000	0.29	0.79	0.50	174.29
			15000 to ≤ 30000	0.48	1.25	0.77	162.00
			> 30000	0.79	1.98	1.19	150.25
	> 4	Divided	≤ 15000	0.33	1.47	1.14	340.78
			15000 to ≤ 30000	0.56	2.34	1.78	321.02
			> 30000	0.92	3.71	2.79	302.15
		Undivided	≤ 15000	0.35	0.89	0.54	153.40
			15000 to ≤ 30000	0.58	1.41	0.83	142.04
			> 30000	0.97	2.24	1.27	131.19

Note: Increase = After – Before

Other Crashes – Table One
Operational Research Modeling

Surrounding Land Use: Urban Location Type: Business or Other
Shoulder: Paved or Other

Intersection Characteristics				Number of Other Crashes			
Speed	Lanes	Median	ADT (vpd)	Before	After	Increase	% Increase
≤ 45 mph	≤ 4	Divided	≤ 15000	0.70	0.70	0.00	0.36
			15000 to ≤ 30000	1.02	1.01	-0.01	-1.18
			> 30000	1.49	1.45	-0.04	-2.69
		Undivided	≤ 15000	0.65	0.70	0.06	8.68
			15000 to ≤ 30000	0.94	1.01	0.07	7.02
			> 30000	1.38	1.45	0.07	5.38
	> 4	Divided	≤ 15000	0.70	0.84	0.14	20.02
			15000 to ≤ 30000	1.02	1.21	0.19	18.18
			> 30000	1.49	1.73	0.24	16.37
		Undivided	≤ 15000	0.65	0.84	0.19	29.97
			15000 to ≤ 30000	0.94	1.21	0.26	27.98
			> 30000	1.38	1.73	0.36	26.02
> 45 mph	≤ 4	Divided	≤ 15000	0.61	0.70	0.10	16.08
			15000 to ≤ 30000	0.88	1.01	0.13	14.30
			> 30000	1.29	1.45	0.16	12.55
		Undivided	≤ 15000	0.56	0.70	0.14	25.71
			15000 to ≤ 30000	0.82	1.01	0.19	23.78
			> 30000	1.19	1.45	0.26	21.89
	> 4	Divided	≤ 15000	0.61	0.84	0.24	38.82
			15000 to ≤ 30000	0.88	1.21	0.32	36.69
			> 30000	1.29	1.73	0.45	34.59
		Undivided	≤ 15000	0.56	0.84	0.28	50.33
			15000 to ≤ 30000	0.82	1.21	0.39	48.02
			> 30000	1.19	1.73	0.54	45.76

Note: Increase = After – Before

Other Crashes – Table Two
Operational Research Modeling

Surrounding Land Use: Rural　　　　Location Type: Business or Other
Shoulder: Paved or Other

Intersection Characteristics				Number of Other Crashes			
Speed	Lanes	Median	ADT (vpd)	Before	After	Increase	% Increase
≤ 45 mph	≤ 4	Divided	≤ 15000	0.51	0.50	-0.01	-1.61
			15000 to ≤ 30000	0.74	0.71	-0.02	-3.12
			> 30000	1.08	1.03	-0.05	-4.60
		Undivided	≤ 15000	0.47	0.50	0.03	6.55
			15000 to ≤ 30000	0.68	0.71	0.03	4.92
			> 30000	0.99	1.03	0.03	3.31
	> 4	Divided	≤ 15000	0.51	0.60	0.09	17.66
			15000 to ≤ 30000	0.74	0.85	0.12	15.86
			> 30000	1.08	1.23	0.15	14.08
		Undivided	≤ 15000	0.47	0.60	0.13	27.42
			15000 to ≤ 30000	0.68	0.85	0.17	25.46
			> 30000	0.99	1.23	0.23	23.54
> 45 mph	≤ 4	Divided	≤ 15000	0.44	0.50	0.06	13.80
			15000 to ≤ 30000	0.64	0.71	0.08	12.06
			> 30000	0.93	1.03	0.10	10.34
		Undivided	≤ 15000	0.40	0.50	0.09	23.24
			15000 to ≤ 30000	0.59	0.71	0.13	21.35
			> 30000	0.86	1.03	0.17	19.49
	> 4	Divided	≤ 15000	0.44	0.60	0.16	36.09
			15000 to ≤ 30000	0.64	0.85	0.22	34.00
			> 30000	0.93	1.23	0.30	31.95
		Undivided	≤ 15000	0.40	0.60	0.19	47.37
			15000 to ≤ 30000	0.59	0.85	0.27	45.12
			> 30000	0.86	1.23	0.37	42.89

Note: Increase = After – Before

APPENDIX E

COMBINED RESULTS

Total Crashes – Table One
Combined Results

Surrounding Land Use: Urban   Location Type: Business   Shoulder: Paved

Intersection Characteristics				Total Number of Crashes			
Speed	Lanes	Median	ADT (vpd)	Before	After	Increase	% Increase
≤ 45 mph	≤ 4	Divided	≤ 15000	3.51	3.14	-0.37	-10.51
			15000 to ≤ 30000	4.24	4.48	0.24	5.78
			> 30000	5.15	6.48	1.33	25.85
		Undivided	≤ 15000	2.48	2.27	-0.21	-8.57
			15000 to ≤ 30000	3.01	3.29	0.27	9.03
			> 30000	3.68	4.82	1.14	30.87
	> 4	Divided	≤ 15000	4.69	3.95	-0.74	-15.79
			15000 to ≤ 30000	5.65	5.65	0.00	0.06
			> 30000	6.86	8.20	1.35	19.69
		Undivided	≤ 15000	3.31	2.86	-0.45	-13.53
			15000 to ≤ 30000	4.01	4.16	0.15	3.67
			> 30000	4.89	6.12	1.23	25.11
> 45 mph	≤ 4	Divided	≤ 15000	2.94	3.14	0.20	6.90
			15000 to ≤ 30000	3.55	4.48	0.93	26.29
			> 30000	4.32	6.48	2.17	50.17
		Undivided	≤ 15000	2.08	2.27	0.19	9.17
			15000 to ≤ 30000	2.53	3.29	0.76	30.11
			> 30000	3.09	4.82	1.73	56.09
	> 4	Divided	≤ 15000	3.92	3.95	0.02	0.60
			15000 to ≤ 30000	4.73	5.65	0.92	19.48
			> 30000	5.74	8.20	2.46	42.84
		Undivided	≤ 15000	2.77	2.86	0.09	3.27
			15000 to ≤ 30000	3.36	4.16	0.80	23.74
			> 30000	4.10	6.12	2.02	49.24

Note: Increase = After – Before

Total Crashes – Table Two
Combined Results

Surrounding Land Use: Urban   Location Type: Business   Shoulder: Other

Intersection Characteristics				Total Number of Crashes			
Speed	Lanes	Median	ADT (vpd)	Before	After	Increase	% Increase
≤ 45 mph	≤ 4	Divided	≤ 15000	3.99	3.50	-0.49	-12.22
			15000 to ≤ 30000	4.83	5.00	0.17	3.49
			> 30000	5.89	7.23	1.34	22.80
		Undivided	≤ 15000	2.83	2.53	-0.30	-10.50
			15000 to ≤ 30000	3.44	3.67	0.22	6.44
			> 30000	4.22	5.38	1.16	27.44
	> 4	Divided	≤ 15000	5.33	4.40	-0.92	-17.34
			15000 to ≤ 30000	6.44	6.31	-0.13	-2.03
			> 30000	7.83	9.16	1.32	16.89
		Undivided	≤ 15000	3.77	3.20	-0.58	-15.28
			15000 to ≤ 30000	4.58	4.64	0.06	1.30
			> 30000	5.60	6.83	1.23	21.93
> 45 mph	≤ 4	Divided	≤ 15000	3.34	3.50	0.16	4.83
			15000 to ≤ 30000	4.05	5.00	0.95	23.52
			> 30000	4.94	7.23	2.30	46.49
		Undivided	≤ 15000	2.37	2.53	0.16	6.84
			15000 to ≤ 30000	2.89	3.67	0.78	27.00
			> 30000	3.54	5.38	1.84	51.95
	> 4	Divided	≤ 15000	4.46	4.40	-0.06	-1.26
			15000 to ≤ 30000	5.39	6.31	0.91	16.96
			> 30000	6.56	9.16	2.59	39.46
		Undivided	≤ 15000	3.16	3.20	0.04	1.15
			15000 to ≤ 30000	3.84	4.64	0.80	20.88
			> 30000	4.70	6.83	2.13	45.41

Note: Increase = After – Before

Total Crashes – Table Three
Combined Results

Surrounding Land Use: Urban   Location Type: Other   Shoulder: Paved

Intersection Characteristics				Total Number of Crashes			
Speed	Lanes	Median	ADT (vpd)	Before	After	Increase	% Increase
≤ 45 mph	≤ 4	Divided	≤ 15000	3.20	2.62	-0.58	-18.07
			15000 to ≤ 30000	3.85	3.75	-0.10	-2.60
			> 30000	4.65	5.42	0.77	16.56
		Undivided	≤ 15000	2.26	1.90	-0.36	-15.90
			15000 to ≤ 30000	2.73	2.75	0.02	0.87
			> 30000	3.31	4.03	0.72	21.80
	> 4	Divided	≤ 15000	4.28	3.30	-0.99	-23.01
			15000 to ≤ 30000	5.14	4.73	-0.41	-8.00
			> 30000	6.20	6.87	0.66	10.69
		Undivided	≤ 15000	3.02	2.40	-0.62	-20.57
			15000 to ≤ 30000	3.64	3.48	-0.15	-4.22
			> 30000	4.41	5.13	0.72	16.27
> 45 mph	≤ 4	Divided	≤ 15000	2.68	2.62	-0.06	-2.08
			15000 to ≤ 30000	3.22	3.75	0.53	16.34
			> 30000	3.90	5.42	1.53	39.15
		Undivided	≤ 15000	1.89	1.90	0.01	0.47
			15000 to ≤ 30000	2.28	2.75	0.47	20.43
			> 30000	2.78	4.03	1.26	45.33
	> 4	Divided	≤ 15000	3.58	3.30	-0.29	-7.98
			15000 to ≤ 30000	4.30	4.73	0.43	9.91
			> 30000	5.19	6.87	1.67	32.17
		Undivided	≤ 15000	2.52	2.40	-0.13	-5.10
			15000 to ≤ 30000	3.04	3.48	0.44	14.37
			> 30000	3.69	5.13	1.43	38.76

Note: Increase = After – Before

Total Crashes – Table Four
Combined Results

Surrounding Land Use: Urban   Location Type: Other   Shoulder: Other

Intersection Characteristics				Total Number of Crashes			
Speed	Lanes	Median	ADT (vpd)	Before	After	Increase	% Increase
≤ 45 mph	≤ 4	Divided	≤ 15000	3.63	2.93	-0.71	-19.46
			15000 to ≤ 30000	4.38	4.18	-0.20	-4.50
			> 30000	5.31	6.05	0.74	13.99
		Undivided	≤ 15000	2.57	2.12	-0.45	-17.49
			15000 to ≤ 30000	3.11	3.07	-0.04	-1.29
			> 30000	3.79	4.50	0.71	18.87
	> 4	Divided	≤ 15000	4.86	3.68	-1.18	-24.26
			15000 to ≤ 30000	5.84	5.28	-0.57	-9.71
			> 30000	7.07	7.66	0.59	8.35
		Undivided	≤ 15000	3.43	2.67	-0.75	-22.01
			15000 to ≤ 30000	4.14	3.89	-0.26	-6.20
			> 30000	5.04	5.72	0.68	13.57
> 45 mph	≤ 4	Divided	≤ 15000	3.04	2.93	-0.11	-3.77
			15000 to ≤ 30000	3.67	4.18	0.51	14.04
			> 30000	4.45	6.05	1.60	36.05
		Undivided	≤ 15000	2.15	2.12	-0.03	-1.46
			15000 to ≤ 30000	2.60	3.07	0.46	17.82
			> 30000	3.17	4.50	1.33	41.80
	> 4	Divided	≤ 15000	4.07	3.68	-0.39	-9.49
			15000 to ≤ 30000	4.89	5.28	0.38	7.83
			> 30000	5.92	7.66	1.74	29.33
		Undivided	≤ 15000	2.87	2.67	-0.20	-6.84
			15000 to ≤ 30000	3.47	3.89	0.42	11.98
			> 30000	4.22	5.72	1.50	35.50

Note: Increase = After – Before

Total Crashes – Table Five
Combined Results

Surrounding Land Use: Rural   Location Type: Business   Shoulder: Paved

Intersection Characteristics				Total Number of Crashes			
Speed	Lanes	Median	ADT (vpd)	Before	After	Increase	% Increase
≤ 45 mph	≤ 4	Divided	≤ 15000	2.99	2.76	-0.22	-7.49
			15000 to ≤ 30000	3.62	3.96	0.34	9.36
			> 30000	4.41	5.74	1.33	30.14
		Undivided	≤ 15000	2.12	2.01	-0.11	-5.40
			15000 to ≤ 30000	2.58	2.91	0.33	12.84
			> 30000	3.16	4.28	1.12	35.47
	> 4	Divided	≤ 15000	3.99	3.48	-0.51	-12.82
			15000 to ≤ 30000	4.82	5.00	0.17	3.62
			> 30000	5.87	7.28	1.41	23.98
		Undivided	≤ 15000	2.83	2.53	-0.29	-10.38
			15000 to ≤ 30000	3.43	3.69	0.26	7.47
			> 30000	4.20	5.45	1.25	29.72
> 45 mph	≤ 4	Divided	≤ 15000	2.50	2.76	0.26	10.47
			15000 to ≤ 30000	3.03	3.96	0.93	30.53
			> 30000	3.70	5.74	2.04	55.25
		Undivided	≤ 15000	1.78	2.01	0.23	12.93
			15000 to ≤ 30000	2.16	2.91	0.75	34.62
			> 30000	2.65	4.28	1.63	61.53
	> 4	Divided	≤ 15000	3.34	3.48	0.14	4.13
			15000 to ≤ 30000	4.04	5.00	0.96	23.70
			> 30000	4.92	7.28	2.36	47.92
		Undivided	≤ 15000	2.37	2.53	0.17	7.01
			15000 to ≤ 30000	2.88	3.69	0.81	28.24
			> 30000	3.52	5.45	1.93	54.71

Note: Increase = After – Before

Total Crashes – Table Six
Combined Results

Surrounding Land Use: Rural Location Type: Business Shoulder: Other

Intersection Characteristics				Total Number of Crashes			
Speed	Lanes	Median	ADT (vpd)	Before	After	Increase	% Increase
≤ 45 mph	≤ 4	Divided	≤ 15000	3.40	3.08	-0.32	-9.37
			15000 to ≤ 30000	4.13	4.42	0.28	6.86
			> 30000	5.05	6.41	1.36	26.83
		Undivided	≤ 15000	2.42	2.24	-0.18	-7.50
			15000 to ≤ 30000	2.95	3.25	0.30	10.03
			> 30000	3.63	4.78	1.15	31.76
	> 4	Divided	≤ 15000	4.54	3.88	-0.66	-14.51
			15000 to ≤ 30000	5.50	5.58	0.07	1.34
			> 30000	6.72	8.12	1.41	20.93
		Undivided	≤ 15000	3.22	2.83	-0.40	-12.30
			15000 to ≤ 30000	3.93	4.12	0.19	4.89
			> 30000	4.81	6.08	1.26	26.27
> 45 mph	≤ 4	Divided	≤ 15000	2.85	3.08	0.23	8.21
			15000 to ≤ 30000	3.46	4.42	0.95	27.52
			> 30000	4.24	6.41	2.17	51.26
		Undivided	≤ 15000	2.03	2.24	0.21	10.39
			15000 to ≤ 30000	2.48	3.25	0.77	31.24
			> 30000	3.04	4.78	1.74	57.07
	> 4	Divided	≤ 15000	3.80	3.88	0.08	2.08
			15000 to ≤ 30000	4.61	5.58	0.97	20.95
			> 30000	5.63	8.12	2.49	44.25
		Undivided	≤ 15000	2.70	2.83	0.13	4.68
			15000 to ≤ 30000	3.29	4.12	0.83	25.13
			> 30000	4.04	6.08	2.04	50.56

Note: Increase = After – Before

Total Crashes – Table Seven
Combined Results

Surrounding Land Use: Rural Location Type: Other Shoulder: Paved

Intersection Characteristics				Total Number of Crashes			
Speed	Lanes	Median	ADT (vpd)	Before	After	Increase	% Increase
≤ 45 mph	≤ 4	Divided	≤ 15000	2.72	2.31	-0.41	-15.11
			15000 to ≤ 30000	3.28	3.31	0.03	0.95
			> 30000	3.98	4.81	0.83	20.84
		Undivided	≤ 15000	1.92	1.68	-0.25	-12.77
			15000 to ≤ 30000	2.33	2.44	0.11	4.66
			> 30000	2.84	3.59	0.75	26.39
	> 4	Divided	≤ 15000	3.64	2.91	-0.73	-20.10
			15000 to ≤ 30000	4.38	4.18	-0.20	-4.49
			> 30000	5.30	6.09	0.79	14.95
		Undivided	≤ 15000	2.57	2.12	-0.45	-17.48
			15000 to ≤ 30000	3.10	3.09	-0.01	-0.46
			> 30000	3.77	4.56	0.79	20.86
> 45 mph	≤ 4	Divided	≤ 15000	2.28	2.31	0.03	1.43
			15000 to ≤ 30000	2.75	3.31	0.56	20.55
			> 30000	3.33	4.81	1.47	44.22
		Undivided	≤ 15000	1.61	1.68	0.07	4.18
			15000 to ≤ 30000	1.95	2.44	0.49	24.92
			> 30000	2.38	3.59	1.21	50.78
	> 4	Divided	≤ 15000	3.05	2.91	-0.14	-4.52
			15000 to ≤ 30000	3.67	4.18	0.52	14.08
			> 30000	4.44	6.09	1.65	37.21
		Undivided	≤ 15000	2.15	2.12	-0.03	-1.43
			15000 to ≤ 30000	2.60	3.09	0.49	18.84
			> 30000	3.16	4.56	1.40	44.20

Note: Increase = After – Before

Total Crashes – Table Eight
Combined Results

Surrounding Land Use: Rural Location Type: Other Shoulder: Other

Intersection Characteristics				Total Number of Crashes			
Speed	Lanes	Median	ADT (vpd)	Before	After	Increase	% Increase
≤ 45 mph	≤ 4	Divided	≤ 15000	3.09	2.58	-0.51	-16.64
			15000 to ≤ 30000	3.74	3.69	-0.04	-1.13
			> 30000	4.54	5.36	0.82	18.04
		Undivided	≤ 15000	2.19	1.87	-0.32	-14.52
			15000 to ≤ 30000	2.66	2.72	0.06	2.29
			> 30000	3.25	4.00	0.75	23.21
	> 4	Divided	≤ 15000	4.13	3.24	-0.89	-21.48
			15000 to ≤ 30000	4.98	4.67	-0.32	-6.38
			> 30000	6.05	6.80	0.75	12.38
		Undivided	≤ 15000	2.92	2.36	-0.56	-19.06
			15000 to ≤ 30000	3.54	3.45	-0.09	-2.63
			> 30000	4.32	5.09	0.77	17.91
> 45 mph	≤ 4	Divided	≤ 15000	2.59	2.58	-0.01	-0.43
			15000 to ≤ 30000	3.13	3.69	0.56	18.03
			> 30000	3.81	5.36	1.56	40.84
		Undivided	≤ 15000	1.83	1.87	0.04	2.07
			15000 to ≤ 30000	2.23	2.72	0.49	22.06
			> 30000	2.72	4.00	1.28	46.94
	> 4	Divided	≤ 15000	3.46	3.24	-0.21	-6.20
			15000 to ≤ 30000	4.17	4.67	0.49	11.79
			> 30000	5.07	6.80	1.73	34.11
		Undivided	≤ 15000	2.45	2.36	-0.08	-3.35
			15000 to ≤ 30000	2.97	3.45	0.48	16.21
			> 30000	3.62	5.09	1.47	40.65

Note: Increase = After – Before

Angle Crashes – Table One
Combined Results

Surrounding Land Use: Urban Location Type: Business Shoulder: Paved
Speed: ≤ 45 mph or > 45 mph

Intersection Characteristics			Number of Angle Crashes			
Lanes	Median	ADT (vpd)	Before	After	Increase	% Increase
≤ 4	Divided	≤ 15000	0.91	0.66	-0.26	-27.92
		15000 to ≤ 30000	1.02	0.85	-0.17	-16.99
		> 30000	1.16	1.10	-0.06	-5.24
	Undivided	≤ 15000	0.57	0.55	-0.02	-3.61
		15000 to ≤ 30000	0.64	0.71	0.07	10.49
		> 30000	0.73	0.91	0.19	25.53
> 4	Divided	≤ 15000	0.91	0.89	-0.03	-2.93
		15000 to ≤ 30000	1.02	1.15	0.12	12.03
		> 30000	1.16	1.48	0.33	28.19
	Undivided	≤ 15000	0.57	0.74	0.17	29.94
		15000 to ≤ 30000	0.64	0.95	0.31	49.27
		> 30000	0.73	1.23	0.51	69.99

Note: Increase = After – Before

Angle Crashes – Table Two
Combined Results

Surrounding Land Use: Urban Location Type: Business Shoulder: Other
Speed: ≤ 45 mph or > 45 mph

Intersection Characteristics			Number of Angle Crashes			
Lanes	Median	ADT (vpd)	Before	After	Increase	% Increase
≤ 4	Divided	≤ 15000	1.35	0.66	-0.69	-51.07
		15000 to ≤ 30000	1.51	0.85	-0.66	-43.70
		> 30000	1.71	1.10	-0.61	-35.80
	Undivided	≤ 15000	0.84	0.55	-0.29	-34.60
		15000 to ≤ 30000	0.94	0.71	-0.24	-25.10
		> 30000	1.07	0.91	-0.16	-14.99
> 4	Divided	≤ 15000	1.35	0.89	-0.46	-34.11
		15000 to ≤ 30000	1.51	1.15	-0.36	-24.01
		> 30000	1.71	1.48	-0.22	-13.15
	Undivided	≤ 15000	0.84	0.74	-0.10	-11.84
		15000 to ≤ 30000	0.94	0.95	0.01	1.18
		> 30000	1.07	1.23	0.16	15.12

Note: Increase = After – Before

Angle Crashes – Table Three
Combined Results

Surrounding Land Use: Urban   Location Type: Other   Shoulder: Paved
Speed: ≤ 45 mph or > 45 mph

Intersection Characteristics			Number of Angle Crashes			
Lanes	Median	ADT (vpd)	Before	After	Increase	% Increase
≤ 4	Divided	≤ 15000	0.91	0.54	-0.38	-41.45
		15000 to ≤ 30000	1.02	0.69	-0.33	-32.50
		> 30000	1.16	0.89	-0.26	-22.87
	Undivided	≤ 15000	0.57	0.44	-0.12	-21.66
		15000 to ≤ 30000	0.64	0.57	-0.06	-10.12
		> 30000	0.73	0.74	0.02	2.23
> 4	Divided	≤ 15000	0.91	0.72	-0.19	-21.01
		15000 to ≤ 30000	1.02	0.93	-0.09	-8.73
		> 30000	1.16	1.21	0.05	4.53
	Undivided	≤ 15000	0.57	0.60	0.03	5.78
		15000 to ≤ 30000	0.64	0.78	0.14	21.66
		> 30000	0.73	1.01	0.28	38.68

Note: Increase = After – Before

Angle Crashes – Table Four
Combined Results

Surrounding Land Use: Urban   Location Type: Other   Shoulder: Other
Speed: ≤ 45 mph or > 45 mph

Intersection Characteristics			Number of Angle Crashes			
Lanes	Median	ADT (vpd)	Before	After	Increase	% Increase
≤ 4	Divided	≤ 15000	1.35	0.54	-0.81	-60.25
		15000 to ≤ 30000	1.51	0.69	-0.82	-54.22
		> 30000	1.71	0.89	-0.82	-47.74
	Undivided	≤ 15000	0.84	0.44	-0.39	-46.85
		15000 to ≤ 30000	0.94	0.57	-0.37	-39.07
		> 30000	1.07	0.74	-0.33	-30.77
> 4	Divided	≤ 15000	1.35	0.72	-0.62	-46.38
		15000 to ≤ 30000	1.51	0.93	-0.57	-38.10
		> 30000	1.71	1.21	-0.50	-29.17
	Undivided	≤ 15000	0.84	0.60	-0.24	-28.22
		15000 to ≤ 30000	0.94	0.78	-0.17	-17.54
		> 30000	1.07	1.01	-0.07	-6.08

Note: Increase = After – Before

Angle Crashes – Table Five
Combined Results

Surrounding Land Use: Rural   Location Type: Business   Shoulder: Paved
Speed: ≤ 45 mph or > 45 mph

Intersection Characteristics			Number of Angle Crashes			
Lanes	Median	ADT (vpd)	Before	After	Increase	% Increase
≤ 4	Divided	≤ 15000	0.72	0.66	-0.06	-7.94
		15000 to ≤ 30000	0.80	0.85	0.05	6.36
		> 30000	0.90	1.10	0.20	21.81
	Undivided	≤ 15000	0.44	0.55	0.10	23.32
		15000 to ≤ 30000	0.50	0.71	0.21	41.81
		> 30000	0.56	0.91	0.35	61.65
> 4	Divided	≤ 15000	0.72	0.89	0.17	23.97
		15000 to ≤ 30000	0.80	1.15	0.35	43.55
		> 30000	0.90	1.48	0.58	64.78
	Undivided	≤ 15000	0.44	0.74	0.29	66.23
		15000 to ≤ 30000	0.50	0.95	0.46	91.59
		> 30000	0.56	1.23	0.67	118.89

Note: Increase = After – Before

Angle Crashes – Table Six
Combined Results

Surrounding Land Use: Rural Location Type: Business Shoulder: Other
Speed: ≤ 45 mph or > 45 mph

Intersection Characteristics			Number of Angle Crashes			
Lanes	Median	ADT (vpd)	Before	After	Increase	% Increase
≤ 4	Divided	≤ 15000	1.05	0.66	-0.40	-37.49
		15000 to ≤ 30000	1.18	0.85	-0.33	-27.85
		> 30000	1.33	1.10	-0.23	-17.44
	Undivided	≤ 15000	0.65	0.55	-0.11	-16.30
		15000 to ≤ 30000	0.73	0.71	-0.03	-3.84
		> 30000	0.83	0.91	0.08	9.51
> 4	Divided	≤ 15000	1.05	0.89	-0.17	-15.81
		15000 to ≤ 30000	1.18	1.15	-0.03	-2.61
		> 30000	1.33	1.48	0.16	11.69
	Undivided	≤ 15000	0.65	0.74	0.08	12.82
		15000 to ≤ 30000	0.73	0.95	0.22	29.91
		> 30000	0.83	1.23	0.40	48.29

Note: Increase = After – Before

Angle Crashes – Table Seven
Combined Results

Surrounding Land Use: Rural Location Type: Other Shoulder: Paved
Speed: ≤ 45 mph or > 45 mph

Intersection Characteristics			Number of Angle Crashes			
Lanes	Median	ADT (vpd)	Before	After	Increase	% Increase
≤ 4	Divided	≤ 15000	0.72	0.54	-0.18	-25.22
		15000 to ≤ 30000	0.80	0.69	-0.11	-13.51
		> 30000	0.90	0.89	-0.01	-0.85
	Undivided	≤ 15000	0.44	0.44	0.00	0.22
		15000 to ≤ 30000	0.50	0.57	0.08	15.36
		> 30000	0.56	0.74	0.18	31.64
> 4	Divided	≤ 15000	0.72	0.72	0.01	0.89
		15000 to ≤ 30000	0.80	0.93	0.14	16.94
		> 30000	0.90	1.21	0.31	34.38
	Undivided	≤ 15000	0.44	0.60	0.16	35.33
		15000 to ≤ 30000	0.50	0.78	0.28	56.14
		> 30000	0.56	1.01	0.44	78.58

Note: Increase = After – Before

Angle Crashes – Table Eight
Combined Results

Surrounding Land Use: Rural   Location Type: Other   Shoulder: Other
Speed: ≤ 45 mph or > 45 mph

Intersection Characteristics			Number of Angle Crashes			
Lanes	Median	ADT (vpd)	Before	After	Increase	% Increase
≤ 4	Divided	≤ 15000	1.05	0.54	-0.52	-49.22
		15000 to ≤ 30000	1.18	0.69	-0.49	-41.33
		> 30000	1.33	0.89	-0.44	-32.80
	Undivided	≤ 15000	0.65	0.44	-0.21	-31.98
		15000 to ≤ 30000	0.73	0.57	-0.16	-21.77
		> 30000	0.83	0.74	-0.09	-10.82
> 4	Divided	≤ 15000	1.05	0.72	-0.33	-31.49
		15000 to ≤ 30000	1.18	0.93	-0.24	-20.66
		> 30000	1.33	1.21	-0.12	-8.92
	Undivided	≤ 15000	0.65	0.60	-0.05	-8.15
		15000 to ≤ 30000	0.73	0.78	0.04	5.88
		> 30000	0.83	1.01	0.17	20.98

Note: Increase = After – Before

Left-turn Crashes – Table One
Combined Results

Surrounding Land Use: Rural or Urban
Shoulder: Paved

Location Type: Business
Speed: ≤ 45 mph or > 45 mph

Intersection Characteristics			Number of Left Turn Crashes			
Lanes	Median	ADT (vpd)	Before	After	Increase	% Increase
≤ 4	Divided	≤ 15000	0.72	0.56	-0.15	-21.59
		15000 to ≤ 30000	0.88	0.70	-0.17	-19.75
		> 30000	1.07	0.88	-0.19	-17.75
	Undivided	≤ 15000	0.52	0.32	-0.21	-39.48
		15000 to ≤ 30000	0.64	0.39	-0.25	-38.41
		> 30000	0.78	0.49	-0.29	-37.23
> 4	Divided	≤ 15000	1.01	0.76	-0.25	-25.14
		15000 to ≤ 30000	1.23	0.95	-0.29	-23.31
		> 30000	1.51	1.19	-0.32	-21.33
	Undivided	≤ 15000	0.73	0.42	-0.31	-42.22
		15000 to ≤ 30000	0.90	0.53	-0.37	-41.15
		> 30000	1.10	0.66	-0.44	-39.96

Note: Increase = After – Before

Left-turn Crashes – Table Two
Combined Results

Surrounding Land Use: Rural or Urban | Location Type: Business
Shoulder: Other | Speed: ≤ 45 mph or > 45 mph

Intersection Characteristics			Number of Left Turn Crashes			
Lanes	Median	ADT (vpd)	Before	After	Increase	% Increase
≤ 4	Divided	≤ 15000	0.86	0.70	-0.16	-18.59
		15000 to ≤ 30000	1.05	0.88	-0.17	-16.61
		> 30000	1.29	1.10	-0.19	-14.46
	Undivided	≤ 15000	0.63	0.39	-0.24	-37.75
		15000 to ≤ 30000	0.77	0.49	-0.28	-36.60
		> 30000	0.95	0.61	-0.33	-35.32
> 4	Divided	≤ 15000	1.21	0.94	-0.27	-22.16
		15000 to ≤ 30000	1.48	1.18	-0.30	-20.19
		> 30000	1.81	1.49	-0.33	-18.05
	Undivided	≤ 15000	0.88	0.53	-0.36	-40.49
		15000 to ≤ 30000	1.08	0.66	-0.43	-39.32
		> 30000	1.33	0.82	-0.51	-38.04

Note: Increase = After – Before

Left-turn Crashes – Table Three
Combined Results

Surrounding Land Use: Rural or Urban
Shoulder: Paved

Location Type: Other
Speed: ≤ 45 mph or > 45 mph

Intersection Characteristics			Number of Left Turn Crashes			
Lanes	Median	ADT (vpd)	Before	After	Increase	% Increase
≤ 4	Divided	≤ 15000	0.56	0.48	-0.08	-14.60
		15000 to ≤ 30000	0.68	0.60	-0.08	-12.17
		> 30000	0.84	0.76	-0.08	-9.55
	Undivided	≤ 15000	0.41	0.27	-0.14	-34.01
		15000 to ≤ 30000	0.50	0.33	-0.16	-32.50
		> 30000	0.61	0.42	-0.19	-30.88
> 4	Divided	≤ 15000	0.79	0.65	-0.15	-18.43
		15000 to ≤ 30000	0.96	0.81	-0.15	-16.04
		> 30000	1.18	1.02	-0.16	-13.44
	Undivided	≤ 15000	0.57	0.36	-0.21	-36.97
		15000 to ≤ 30000	0.70	0.45	-0.25	-35.49
		> 30000	0.85	0.56	-0.29	-33.86

Note: Increase = After – Before

Left-turn Crashes – Table Four
Combined Results

Surrounding Land Use: Rural or Urban — Location Type: Other
Shoulder: Other — Speed: ≤ 45 mph or > 45 mph

Intersection Characteristics			Number of Left Turn Crashes			
Lanes	Median	ADT (vpd)	Before	After	Increase	% Increase
≤ 4	Divided	≤ 15000	0.67	0.60	-0.07	-10.65
		15000 to ≤ 30000	0.82	0.75	-0.07	-8.04
		> 30000	1.00	0.95	-0.05	-5.21
	Undivided	≤ 15000	0.49	0.33	-0.15	-31.59
		15000 to ≤ 30000	0.60	0.42	-0.18	-29.98
		> 30000	0.73	0.52	-0.21	-28.23
> 4	Divided	≤ 15000	0.94	0.81	-0.14	-14.54
		15000 to ≤ 30000	1.15	1.02	-0.14	-11.95
		> 30000	1.41	1.28	-0.13	-9.15
	Undivided	≤ 15000	0.69	0.45	-0.24	-34.59
		15000 to ≤ 30000	0.84	0.56	-0.28	-32.98
		> 30000	1.03	0.71	-0.32	-31.23

Note: Increase = After – Before

Rear-end Crashes – Table One
Combined Results

Surrounding Land Use: Rural or Urban — Location Type: Business
Shoulder: Paved

Intersection Characteristics				Number of Rear-end Crashes			
Speed	Lanes	Median	ADT (vpd)	Before	After	Increase	% Increase
≤ 45 mph	≤ 4	Divided	≤ 15000	0.50	0.80	0.30	60.84
			15000 to ≤ 30000	0.85	1.38	0.53	63.20
			> 30000	1.44	2.40	0.96	66.90
		Undivided	≤ 15000	0.55	0.55	0.00	0.28
			15000 to ≤ 30000	0.94	0.96	0.03	2.94
			> 30000	1.59	1.69	0.10	6.49
	> 4	Divided	≤ 15000	0.68	0.93	0.26	38.16
			15000 to ≤ 30000	1.15	1.62	0.46	40.29
			> 30000	1.96	2.82	0.86	43.57
		Undivided	≤ 15000	0.75	0.65	-0.10	-13.91
			15000 to ≤ 30000	1.28	1.13	-0.15	-11.56
			> 30000	2.19	2.00	-0.18	-8.45
> 45 mph	≤ 4	Divided	≤ 15000	0.33	0.78	0.45	134.36
			15000 to ≤ 30000	0.57	1.34	0.77	135.93
			> 30000	0.96	2.31	1.34	139.34
		Undivided	≤ 15000	0.37	0.53	0.16	44.50
			15000 to ≤ 30000	0.63	0.92	0.29	47.14
			> 30000	1.06	1.61	0.54	51.00
	> 4	Divided	≤ 15000	0.45	0.91	0.46	101.97
			15000 to ≤ 30000	0.77	1.56	0.80	103.47
			> 30000	1.31	2.70	1.39	106.56
		Undivided	≤ 15000	0.50	0.62	0.12	24.45
			15000 to ≤ 30000	0.85	1.08	0.23	26.82
			> 30000	1.45	1.89	0.44	30.24

Note: Increase = After – Before

Rear-end Crashes – Table Two
Combined Results

Surrounding Land Use: Rural or Urban
Location Type: Business
Shoulder: Other

Intersection Characteristics				Number of Rear-end Crashes			
Speed	Lanes	Median	ADT (vpd)	Before	After	Increase	% Increase
≤ 45 mph	≤ 4	Divided	≤ 15000	0.50	1.09	0.59	118.94
			15000 to ≤ 30000	0.85	1.88	1.04	122.48
			> 30000	1.44	3.28	1.84	127.86
		Undivided	≤ 15000	0.55	0.75	0.20	36.82
			15000 to ≤ 30000	0.94	1.32	0.38	40.66
			> 30000	1.59	2.32	0.73	45.72
	> 4	Divided	≤ 15000	0.68	1.27	0.60	88.18
			15000 to ≤ 30000	1.15	2.21	1.05	91.37
			> 30000	1.96	3.85	1.89	96.13
		Undivided	≤ 15000	0.75	0.88	0.13	17.52
			15000 to ≤ 30000	1.28	1.55	0.27	20.91
			> 30000	2.19	2.74	0.55	25.34
> 45 mph	≤ 4	Divided	≤ 15000	0.33	1.07	0.73	218.48
			15000 to ≤ 30000	0.57	1.82	1.26	221.08
			> 30000	0.96	3.14	2.18	226.20
		Undivided	≤ 15000	0.37	0.73	0.36	96.80
			15000 to ≤ 30000	0.63	1.26	0.63	100.70
			> 30000	1.06	2.19	1.13	106.27
	> 4	Divided	≤ 15000	0.45	1.24	0.79	174.63
			15000 to ≤ 30000	0.77	2.13	1.36	177.07
			> 30000	1.31	3.68	2.38	181.70
		Undivided	≤ 15000	0.50	0.85	0.35	69.59
			15000 to ≤ 30000	0.85	1.47	0.62	73.08
			> 30000	1.45	2.58	1.13	78.02

Note: Increase = After – Before

Rear-end Crashes – Table Three
Combined Results

Surrounding Land Use: Rural or Urban   Location Type: Other
Shoulder: Paved

Intersection Characteristics				Number of Rear-end Crashes			
Speed	Lanes	Median	ADT (vpd)	Before	After	Increase	% Increase
≤ 45 mph	≤ 4	Divided	≤ 15000	0.37	0.80	0.43	116.94
			15000 to ≤ 30000	0.63	1.38	0.75	119.87
			> 30000	1.07	2.40	1.33	124.57
		Undivided	≤ 15000	0.41	0.55	0.14	34.90
			15000 to ≤ 30000	0.70	0.96	0.27	38.32
			> 30000	1.19	1.69	0.51	42.91
	> 4	Divided	≤ 15000	0.50	0.93	0.43	85.35
			15000 to ≤ 30000	0.86	1.62	0.76	87.98
			> 30000	1.47	2.82	1.35	92.14
		Undivided	≤ 15000	0.56	0.65	0.09	15.19
			15000 to ≤ 30000	0.96	1.13	0.17	18.19
			> 30000	1.64	2.00	0.36	22.19
> 45 mph	≤ 4	Divided	≤ 15000	0.25	0.78	0.54	217.12
			15000 to ≤ 30000	0.42	1.34	0.92	218.86
			> 30000	0.71	2.31	1.59	223.08
		Undivided	≤ 15000	0.27	0.53	0.26	95.00
			15000 to ≤ 30000	0.46	0.92	0.46	98.35
			> 30000	0.79	1.61	0.82	103.30
	> 4	Divided	≤ 15000	0.34	0.91	0.58	171.85
			15000 to ≤ 30000	0.57	1.56	0.99	173.54
			> 30000	0.97	2.70	1.73	177.34
		Undivided	≤ 15000	0.37	0.62	0.25	67.06
			15000 to ≤ 30000	0.64	1.08	0.44	70.03
			> 30000	1.08	1.89	0.81	74.41

Note: Increase = After – Before

Rear-end Crashes – Table Four
Combined Results

Surrounding Land Use: Rural or Urban   Location Type: Other
Shoulder: Other

Intersection Characteristics				Number of Rear-end Crashes			
Speed	Lanes	Median	ADT (vpd)	Before	After	Increase	% Increase
≤ 45 mph	≤ 4	Divided	≤ 15000	0.37	1.09	0.72	195.29
			15000 to ≤ 30000	0.63	1.88	1.25	199.74
			> 30000	1.07	3.28	2.21	206.60
		Undivided	≤ 15000	0.41	0.75	0.34	84.06
			15000 to ≤ 30000	0.70	1.32	0.62	88.99
			> 30000	1.19	2.32	1.13	95.55
	> 4	Divided	≤ 15000	0.50	1.27	0.77	152.46
			15000 to ≤ 30000	0.86	2.21	1.35	156.42
			> 30000	1.47	3.85	2.39	162.48
		Undivided	≤ 15000	0.56	0.88	0.32	57.25
			15000 to ≤ 30000	0.96	1.55	0.59	61.58
			> 30000	1.64	2.74	1.10	67.30
> 45 mph	≤ 4	Divided	≤ 15000	0.25	1.07	0.82	330.95
			15000 to ≤ 30000	0.42	1.82	1.40	333.94
			> 30000	0.71	3.14	2.43	340.33
		Undivided	≤ 15000	0.27	0.73	0.45	165.59
			15000 to ≤ 30000	0.46	1.26	0.79	170.55
			> 30000	0.79	2.19	1.40	177.72
	> 4	Divided	≤ 15000	0.34	1.24	0.91	269.66
			15000 to ≤ 30000	0.57	2.13	1.56	272.48
			> 30000	0.97	3.68	2.71	278.23
		Undivided	≤ 15000	0.37	0.85	0.48	127.67
			15000 to ≤ 30000	0.64	1.47	0.84	132.06
			> 30000	1.08	2.58	1.50	138.39

Note: Increase = After – Before

Other Crashes – Table One
Combined Results

Surrounding Land Use: Urban — Location Type: Business or Other
Shoulder: Paved or Other

Intersection Characteristics				Number of Other Crashes			
Speed	Lanes	Median	ADT (vpd)	Before	After	Increase	% Increase
≤ 45 mph	≤ 4	Divided	≤ 15000	0.76	0.63	-0.14	-17.78
			15000 to ≤ 30000	1.09	0.97	-0.12	-10.99
			> 30000	1.55	1.51	-0.05	-3.07
		Undivided	≤ 15000	0.64	0.63	-0.01	-1.62
			15000 to ≤ 30000	0.91	0.97	0.06	6.30
			> 30000	1.30	1.51	0.20	15.52
	> 4	Divided	≤ 15000	0.76	0.77	0.01	1.00
			15000 to ≤ 30000	1.09	1.19	0.10	9.61
			> 30000	1.55	1.86	0.31	19.64
		Undivided	≤ 15000	0.64	0.77	0.13	20.86
			15000 to ≤ 30000	0.91	1.19	0.28	30.89
			> 30000	1.30	1.86	0.56	42.59
> 45 mph	≤ 4	Divided	≤ 15000	0.57	0.63	0.06	10.13
			15000 to ≤ 30000	0.82	0.97	0.15	18.87
			> 30000	1.17	1.51	0.34	29.05
		Undivided	≤ 15000	0.48	0.63	0.14	29.96
			15000 to ≤ 30000	0.69	0.97	0.28	39.99
			> 30000	0.99	1.51	0.51	51.69
	> 4	Divided	≤ 15000	0.57	0.77	0.20	35.29
			15000 to ≤ 30000	0.82	1.19	0.38	46.38
			> 30000	1.17	1.86	0.69	59.29
		Undivided	≤ 15000	0.48	0.77	0.29	59.65
			15000 to ≤ 30000	0.69	1.19	0.50	72.38
			> 30000	0.99	1.86	0.87	87.23

Note: Increase = After – Before

Other Crashes – Table Two
Combined Results

Surrounding Land Use: Rural Location Type: Business or Other
Shoulder: Paved or Other

Intersection Characteristics				Number of Other Crashes			
Speed	Lanes	Median	ADT (vpd)	Before	After	Increase	% Increase
≤ 45 mph	≤ 4	Divided	≤ 15000	0.55	0.45	-0.09	-17.10
			15000 to ≤ 30000	0.78	0.70	-0.08	-10.10
			> 30000	1.11	1.09	-0.02	-1.94
		Undivided	≤ 15000	0.46	0.45	0.00	-0.87
			15000 to ≤ 30000	0.65	0.70	0.05	7.28
			> 30000	0.93	1.09	0.16	16.79
	> 4	Divided	≤ 15000	0.55	0.56	0.01	1.91
			15000 to ≤ 30000	0.78	0.86	0.08	10.77
			> 30000	1.11	1.35	0.23	21.11
		Undivided	≤ 15000	0.46	0.56	0.10	21.86
			15000 to ≤ 30000	0.65	0.86	0.21	32.19
			> 30000	0.93	1.35	0.41	44.24
> 45 mph	≤ 4	Divided	≤ 15000	0.41	0.45	0.04	10.92
			15000 to ≤ 30000	0.58	0.70	0.12	19.93
			> 30000	0.84	1.09	0.25	30.42
		Undivided	≤ 15000	0.35	0.45	0.11	30.79
			15000 to ≤ 30000	0.50	0.70	0.20	41.14
			> 30000	0.71	1.09	0.38	53.19
	> 4	Divided	≤ 15000	0.41	0.56	0.15	36.36
			15000 to ≤ 30000	0.58	0.86	0.28	47.77
			> 30000	0.84	1.35	0.51	61.09
		Undivided	≤ 15000	0.35	0.56	0.21	60.79
			15000 to ≤ 30000	0.50	0.86	0.37	73.92
			> 30000	0.71	1.35	0.64	89.21

Note: Increase = After – Before