2017

Examining the safety performance of intersections on arterial roadways and near service ramp terminals

Timothy P. Barrette
Iowa State University

Examining the safety performance of intersections on arterial roadways and near service ramp terminals

by

Timothy P. Barrette

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Civil Engineering (Transportation Engineering)

Program of Study Committee:
Peter T. Savolainen, Major Professor
R. Christopher Williams
Omar Smadi
Anuj Sharma
Alicia Carriquiry

Iowa State University

Ames, Iowa

2017

DEDICATION

This dissertation is dedicated to my friends and family in gratitude for their support throughout my education. I can never thank my parents, Tim and Sally, enough for encouraging me to pursue this degree. I am also very thankful to Uncle Jerry and Aunt Nan for housing me while when I first went back to school. Finally, I am extremely grateful to have had my wife, Emira, beside me and enduring the same challenges throughout this process.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

The completion of this document would not have been possible without the hard work of a large and talented team of current and former student researchers. Much of the data collected for the Michigan studies was done with extensive effort from Emira Rista (currently with the Texas A&M Transportation Institute), Dr. Brendan Russo (currently of Northern Arizona University), Raha Hamezie.

I am extremely grateful to Patricia Thompson, Iftin Thompson, Akinfolarin Abatan, Sawyer Reynolds, Eric Ure, Pedro Paiva de Lima, and Ashirwad Barnwal for their efforts in the collection of data used for the Iowa study.

Finally, I am very grateful to have had the chance to work for Dr. Peter Savolainen for the past four years and extremely appreciative of the resulting opportunities.

ABSTRACT

Recent transportation bills have required each state to have a Highway Safety Improvement Program which emphasizes a data-driven approach to improving highway safety. This data-driven paradigm, coupled with the fact that approximately 40 percent of traffic crashes in the U.S. occur at intersections, has led to substantial research focusing on intersections. This study focuses on three areas of intersection safety: vehicular crashes, pedestrian and cyclist crashes, and crashes near ramp-terminal intersections. The impact of geometric characteristics on vehicular crashes at intersections using five years of crash data at an aggregate and disaggregate level. The within sample predictive ability of negative binomial models estimated using aggregate crash data (with empirical Bayes methodology) was compared to that of a disaggregate model estimated using a site-specific random effects negative binomial framework. Pedestrian and cyclist crashes are often difficult to model on a large scale as exposure information is typically not collected or maintained by road agencies. To this end, the characteristics affecting pedestrian and cyclist crashes at intersections have been examined using census tract-level commuter information from the American Community Survey in lieu of observed pedestrian and cyclist volume. Finally, ramp terminal intersections provide important points of connection between restricted access roadways (such as interstates) and adjacent land. The safety performance along the corridors adjacent to the ramp terminal intersections is directly related to the proximity between the ramp terminal and access points such as driveways and intersections. This study explores the effect of ramp terminal and access point proximity on corridor safety and provides a framework for road agencies to evaluate corridor-level safety implications based on the proximity between ramp terminals and access points, the volume of the crossroad, and the volume of the access point.

CHAPTER 1. INTRODUCTION

1.1 Statement of the Problem

The Fixing America's Surface Transportation (FAST) Act has continued to build upon the foundation laid by the enactment of the Moving Ahead for Progress in the 21st Century Act (MAP-21), which required that each state have in place a Highway Safety Improvement Program (HSIP). The purpose of the HSIP is to ensure every state "emphasizes a data-driven, strategic approach to improving highway safety on all public roads that focuses on performance". This emphasis toward data-driven safety practices has resulted in significant research focusing on better understanding how various geometric characteristics and traffic conditions affect the frequency, type, and severity of crashes on specific roadway facility types such as road segments, ramps, and intersections. Gaining an improved understanding of the relationships between roadway characteristics and crashes will enable transportation safety professionals to develop appropriate and proactive policies and countermeasures that reduce potential for traffic crashes and the resultant injuries and fatalities.

The first edition of the *Highway Safety Manual* (*HSM*) was developed by the American Association of State Highway and Transportation Officials (AASHTO) to provide guidance as to best practices that allow for prediction of the safety performance of road facilities with specific site conditions (AASHTO 2010). The HSM also provides a series of commonly used tool that facilitate the understanding of the relationships between roadway characteristics and crashes. Part C of the HSM provides a series of crash prediction models, commonly referred to as *safety performance functions (SPFs)*, which can be utilized to estimate the frequency of traffic crashes on specific road facilities as a function of traffic volumes, roadway geometry, type of traffic control, and other factors. SPFs can be useful for estimating the safety impacts of site-specific design alternatives through the implementation of an empirical-Bayes analysis or for prioritizing candidate locations for safety improvements on a network basis. Various software programs and support tools have been designed to utilize SPFs, such as *SafetyAnalyst* (Lu et al. 2012) and the *Interactive Highway Safety Design Model* (*IHSDM*) (Lubliner et al. 2015).

The HSM presents a series of SPFs for intersections and road segments of specific road types, such as rural two-lane highways, rural multi-lane highways, and urban/suburban arterial roadways. Various states have conducted research pertaining to the utilization of the SPFs from the HSM, as well as to the development of jurisdiction-specific SPFs that have been estimated using local data or calibrated to local conditions. These studies have been conducted in states including Colorado, Florida, Georgia, Illinois, Kansas, North Carolina, Oregon, Utah, and Virginia (Garber et al. 2010A; Persaud and Lyon 2009; Garber et al. 2010B; Tegge et al. 2010; Dixon et al 2012; Wang et al. 2011; Srinivasan, R. et al. 2011; Brimley et al. 2012; Bornheimer et al. 2012; Lu et al. 2012; Lubliner et al. 2012; Srinivasan, S. et al. 2011, Alluri et al. 2012A).

Collectively, these studies have illustrated several important shortcomings of the base SPFs from the HSM. Among the most important of these concerns is that the accuracy of these SPFs has been found to vary considerably from state to state, a result that may be reflective of differences in geography, design practices, driver behavior, differences in crash reporting requirements, or other factors.

Intersections are one of the facility types of greatest concerns as they have been shown to contribute to nearly 25 percent of annual fatalities in the United States and roughly half of all traffic-related injuries (Subramanian and Lombardo 2007). The HSM presents SPFs for intersections located on urban and suburban arterials, which were developed based on data from Minnesota and North Carolina and subsequently validated on data from Florida (Harwood et al.

2008). Given concerns as to the transferability of the base SPFs from the HSM for this facility type, this study involved the estimation of SPFs specific to urban/suburban intersections in the state of Michigan. The accuracy of the Michigan-specific SPFs is compared to results from the base SPFs from the HSM, both before and after calibration using local crash data.

Although the HSM provides crash prediction methodologies for a vast array of roadway facilities and crash types, there are several areas that would benefit from additional research. One area that is somewhat under researched is the development of crash prediction models for pedestrian- and bicyclist-involved crashes. In order to utilize the SPFs for pedestrian and bicycle crashes included in the HSM, road agencies must have access to traffic counts of pedestrian and bicyclist activity, or the means to estimate those volumes. In many cases, particularly in the case of large-scale (i.e., statewide) studies, such data are generally not available. In order to facilitate the development of SPFs at a statewide level in a manner that is easily replicable, this study utilized data from the American Community Survey (ACS) to investigate the use of census-tract characteristics such as the density of walking or cycling commuters as pseudo-exposure measures. The ACS is publicly available data and road agencies with sufficient data from the survey can utilize the approach presented in this study to predict pedestrian and cyclist crashes in lieu of detailed count data. This study makes use of data from Michigan to demonstrate an approach for the development of crash prediction models for pedestrians and cyclists.

Finally, this study examines the safety performance of road segments immediately downstream of the crossroad ramp terminal at service interchanges. This study examines how the rate of crashes varies based upon the proximity to the nearest access points (i.e., intersections or driveways) immediately downstream of the freeway exit ramp. Data from all service interchanges in the state of Iowa are used as a part of this study, which has two primary objectives: 1) to develop guidelines for appropriate access point location on a crossroad location relative to ramp bifurcation points, and 2) to determine a minimum crossroad volume threshold at which point the proximity guidance is applicable (i.e., determining whether is there a minimum volume threshold below which the proximity of the first access point has negligible impact).

The remainder of this document is dedicated to addressing the three high-level issues pertaining to intersection safety described above. A brief summary of the subsequent chapters if provided here:

- Chapter 2 details the development SPFs for vehicular crashes on arterial roadways. These SPFs were estimated as volume-only models, volume and regional indicator models, as well as detailed geometric models. The variety of model complexity was intended to facilitate usage in circumstances when all data characteristics may not be known.
- Chapter 3 details the development of pedestrian, cyclist, and combined non-motorist SPFs. These models improve upon existing large-scale pedestrian and cyclist crash prediction models by incorporating census-tract level data to provide a measure of exposure which has not been utilized in the extant literature.
- Chapter 4 examines the effect of access point proximity relative to freeway ramp-terminal intersections. This research provides guidance to road agencies as to the general relationship that exists between access points location relative to ramp-terminal intersections and the resulting crash rate.
- The appendix includes an additional paper that was published during the course of the dissertation research, but on a topic outside the general focus of this

dissertation. This paper investigates the degree of injury severity sustained by crash-involved motorcyclists in consideration of changes to Michigan's motorcycle helmet use law.

CHAPTER 2. SAFETY PERFORMANCE FUNCTIONS FOR VEHICULAR CRASHES

2.1 Study Objectives

This study involves the estimation of SPFs for four types of intersections:

1. Four-leg signalized (4SG);
2. Four-leg stop-controlled (4ST);
3. Three-leg signalized (3SG); and
4. Three-leg stop-controlled (3ST).

In addition to providing important tools for use in network screening and evaluation, this study also examines several important questions of interest to the broader safety research community. While a considerable amount of research has documented the calibration or estimation of jurisdiction-specific SPFs, the extant literature offers minimal guidance as to several important questions of concern to state and local road agencies. For example, the HSM provides limited guidance as to when the development of specific SPFs is desirable and/or feasible. To this end, several additional objectives are proposed as a part of this study:

- Explore the implications of developing SPFs utilizing aggregate vs disaggregate data. More specifically, comparing the predictive capabilities of SPFs estimated on a data set where one observation represents five years of crashes at a specific site, versus where one observation represents one year. To account for repeated observations, a random effects model was estimated on the disaggregate data. The predictive capabilities were then analyzed naively (without site-specific intercepts) and then precisely (using site-specific intercept terms on the disaggregate data and empirical-Bayes calibration on the aggregate data).

- Examine the granularity of safety performance across various geographic regions (e.g., statewide, regional, county, MPO, city). Typically, SPFs are estimated using a sample of statewide data. However, as prior research has shown differences in safety performance across states, it is also reasonable to expect significant variability in safety trends within a single state.
- Investigate the impact of including regional indicator variables in SPFs. Typically, SPFs are estimated using a sample of statewide data. However, as prior research has shown differences in safety performance across states, it is also reasonable to expect significant variability in safety trends within a single state.

To accomplish these goals as well as make significant contributions to the research literature, an in-depth review of extant literature, data collection, and data analysis were conducted. These items are all thoroughly outlined in the remainder of this document.

2.2 Literature Review

Given the current emphases on data-driven strategic approaches for safety analysis, a priority area at the national level has been the identification of high-risk intersections and road segments. Site identification is a critical component of a safety improvement program and the effective identification of sites that are candidates for improvements can be costly (Hauer et al. 2002). Historically, a variety of methods have been used to identify and prioritize candidate sites for safety treatments. These have largely included simple methods such as the ranking of sites based upon system-wide crash frequency or crash rate data. There are several drawbacks to such approaches. For example, considering only crash frequency tends to ignore sites with low traffic volumes while using crash rates tends to disproportionately prioritize very low volume sites (Persaud 2001). The use of crash rates also implicitly assumes a linear relationship between crashes and traffic volume, which is not necessarily well supported by safety research (AASHTO 2015B). However, due to the minimal data requirements, these methods are still widely used by DOTs in site screening and the identification of crash hot spots (Alluri 2008).

A bigger concern is that, given the random nature of crashes on a location-by-location basis, short-term trends in crash frequency or rate are not necessarily good predictors of long-term crash frequency (Alluri). This concern relates largely to a phenomenon called regression-to-the-mean (RTM). In practical terms, RTM is reflected by the fact that roadway locations that experience particularly high short-term (e.g., one year) crash frequencies are likely to decrease closer to the average of similar sites (i.e., regress to the mean) over the long term (Hauer 1997). To address such concerns, short-term site-specific crash counts can be combined with estimates from predictive regression models to develop more accurate estimates of long-term (i.e., future) safety performance. An important tool in this process is the American Association of State Highway and Transportation Officials (AASHTO) Highway Safety Manual (HSM) (AASHTO 2010). Part C of the HSM provides a series of predictive models, referred to as safety performance functions (SPFs), which can be utilized to estimate the frequency of traffic crashes on specific road facilities as a function of traffic volumes, roadway geometry, type of traffic control, and other factors.

SPFs establish a basis for evaluating roadway safety in consideration of the effects of traffic volume (AADT) roadway geometry, and other factors. SPFs for intersections take the following general form given in Equation 1:

$$N_{spf} = exp(\beta_0)AADT_{major}^{\beta_1}AADT_{minor}^{\beta_2}, \tag{1}$$

where:
N_{spf} = predicted average crash frequency for a site with base conditions;
$AADT_{major}$ = annual average daily traffic (AADT) for the major road;
$AADT_{minor}$ = annual average daily traffic (AADT) for the minor road; and
$\beta_0, \beta_1, \beta_2$ = estimated parameters.

Although the HSM provides default SPF models, it is noteworthy that these models were developed using data from only a few states. This makes the transferability of the SPFs a critical issue that needs to be handled by state agencies and DOTs when they attempt to implement these models. While these SPFs can be directly applied, the HSM recommends that the equations are either calibrated using local (i.e., state or regional) data or that jurisdiction-specific SPFs are developed (AASHTO 2010). The calibrated model must sufficiently capture local road and traffic features (Chen et al. 2012). Calibration of the SPFs is relatively straightforward, requiring the estimation of a calibration factor, C, as shown in Equation 2:

$$N_{predicted} = N_{spf} \times C\ , \tag{2}$$

where:
$N_{predicted}$ = predicted annual average crash frequency for a specific site;
N_{spf} = predicted average crash frequency for a site with base conditions; and
C = calibration factor to adjust SPF for local conditions.

This calibration factor is simply equal to the ratio of the number of observed crashes within the jurisdiction to the predicted number of crashes as estimated by the SPF. While calibration generally results in improved goodness-of-fit, research has shown that the suggested sample sizes for sites (30-50) and crashes (100 per year) in the HSM do not necessarily minimize predictive error in calibration (Shin et al. 2014).

In addition to calibration for local factors, it is also important to note that the SPFs from the HSM are estimated for “base” conditions. At locations where base conditions are not met, the

SPFs are multiplied by crash modification factors (CMFs), which adjust the SPF for non-base conditions as shown in Equation 3:

$$N_{predicted} = N_{sp} \times C \times CMF_i \,, \tag{3}$$

where:
$N_{predicted}$ = predicted annual average crash frequency for a specific site;
N_{spf} = predicted average crash frequency for a site with base conditions;
C = calibration factor to adjust SPF for local conditions; and
CMF_i = crash modification factor for condition i.

These CMFs allow for crash estimates that distinguish between sites with various geometric or traffic control features. For example, the HSM provides a series of CMFs in Chapter 12 specific to intersections on urban and suburban arterials. Chapter 14 provides a catalog of various intersection CMFs based on prior empirical research. In addition, the Federal Highway Administration (FHWA) maintains the Crash Modification Factor (CMF) Clearinghouse (FHWA 2015), a web-based database of CMFs that provides supporting documentation to assist users in estimating the impacts of various safety countermeasures.

2.2.1 Summary of State Efforts in SPF Calibration and Development

A recent study summarized the results of a nation-wide survey that was employed to assess the current status of safety analysis at state departments of transportation (Alluri and Ogle 2012B). The results of this survey demonstrated that most states experienced data-related issues that inhibited their ability to effectively conduct safety analyses. A Florida study cited the data requirements of the HSM were challenging as many of the factors were not available in the state's roadway characteristics inventory database (Alluri et al. 2014). Similar results were found in Pennsylvania where several variables suggested in the HSM could not be included in SPFs due to lack of available data (Donnell et al. 2014). Several other studies have also identified data availability and completeness as hurdles in meeting the input requirements of the HSM and other related tools such as SafetyAnalyst (Alluri et al. 2014; Donnell et al. 2014). A study in Georgia found that data quality and availability significantly affect the quality and reliability of SPFs (Alluri and Ogle 2012A) while research in Kansas noted that the scarcity of intersection data did not allow for the development of separate models for 3-leg and 4-leg stop-controlled intersections (Lubliner et al. 2014).

Specific areas of concern included lack of sufficient data on traffic volumes and roadway characteristics, as well as a lack of geo-referenced spatial data (Alluri and Ogle 2012B). In most states, traffic data is generally available for higher classes of roadways (e.g., interstates, state routes, etc.), but is limited for local and low volume roads (Alluri and Ogle 2012B). Research in Colorado found that volume data for side-streets were not generally available for more than one or two years, and in many cases the count data did not coincide with the study period (Persaud and Lyon 2009). Thus, it was necessary to normalize available side-street AADT data over the study period using growth rates derived from the mainline AADT volumes (Persaud and Lyon 2009).

International studies (Sacchi et al. 2012; Cafiso et al. 2012; Cunto et al. 2015; Persaud and Nguyen 1998; Giuffre et al. 2014; Martinelli et al. 2009) also show that sampling of sites is often hindered by the availability of data. Studies in Brazil (Cunto et al. 2015) and Italy (Giuffre et al. 2014; Martinelli et al. 2009) found the need for manually collected data on traffic volumes, roadway geometry, and functional characteristics limited the number of sites that could feasibly be included in SPF estimation.

Despite these limitations, Table 1 shows a significant number of recent state-level efforts aimed at either calibrating the HSM SPFs or developing state-specific SPFs using local data. The table summarizes recent studies, including details of the types of intersections that were considered as a part of each study, the number of sites that were included by type, and the number of years of data that were used for model calibration of estimation.

When examining SPF calibration for local conditions, there is significant variability in terms of whether the base models from the HSM over- or under-predict crashes within specific states. Research in Kansas found a calibration factor of 0.21, indicating that crashes were significantly over-predicted at unsignalized three-leg and four-leg intersections in the state (Lubliner et al. 2014). However, these studies note that the calibration factors were developed using a small sample dataset and, as such, they should be used with caution. Calibration factors for urban intersections in Maryland ranged from 0.1562 for three-leg stop controlled intersections to 0.4747 for four-leg signalized intersections (Shin et al. 2014). Research in Oregon (Monsere et al. 2011) and North Carolina (Srinivasan and Carter 2011) also tended to show significantly lower crashes than would be predicted by the base models from the HSM. Statewide HSM model calibration in Missouri generally showed calibration factors less than 1.0, suggesting that Missouri facilities experienced fewer crashes than the national average (Sun et al. 2014). However, the converse was true for urban three-leg and four-leg signalized intersections, where calibration factors of 3.03 and 4.91 were observed, respectively. The magnitude of these calibration factors was attributed to differences in crash definitions between Missouri and the states used as the basis for the HSM.

Table 1. Summary of studies involving calibration or development of specific SPFs

State/ Country	Site Type(s)	No. of Sites	No. of Years	Calibrated HSM SPFs	Jurisdiction Specific SPFs
AB	4SG	99	3-7	No	Yes
AB; ON	4SG	515; 1629	6	Yes	Yes
Brazil	4SG; 4ST	353; 132	6	No	Yes
Brazil	4SG; 4ST	89; 92	3	Yes	No
BC	SG	98	9	No	Yes
BC	SG	51	3	No	Yes
CA; ON	4ST; 3SG/4SG	2202; >20	-	No	Yes
CA	3ST; 4ST	378; 264	10	No	Yes
CA	3ST, 4ST	1381, 907	10	No	No
FL	4SG	519	6	No	Yes
FL	4SG	177	6	No	Yes
FL	3ST, 4ST, 4SG	31-321; 58; 34-43; 21-459	3	No	Yes
ITA	4ST (one-way)	92	7	No	Yes
MD	3ST, 4ST, 4SG	152-162; 26-167; 10-115; 35-244	3	Yes	No
MO	3ST, 4ST	35-70; 25-70	1	Yes	No
OH	3ST, 4ST, 4SG	50-200; 50-200; 125-250; 50-200	3	Yes	No
ON	3SG; 4SG	40; 230	6	No	Yes
ON	3ST; 3SG; 4ST; 4SG	117; 250; 59; 868	6	No	Yes
ON	3SG, 4SG	306, 1410	5	Yes	Yes
ON	3SG	59	6	No	Yes
ON	3SG; 4SG	137; 1691	6	Yes	Yes
OR	3ST; 4SG	202; 298	3	Yes	Yes
PA	3ST; 4ST; 3SG; 4SG	414; 86; 45; 105	8	No	Yes
SK	3ST; 4ST; 3SG/4SG	123; 121; 143	5	Yes	Yes
South Korea	3SG; 4SG	247; 201	2	No	Yes
VA	3ST; 3SG; 4ST; 4SG	5367-8411; 183-836; 1239-1570; 182-568	6	No	Yes
VA	4SG	35	4	No	Yes
VA	4SG	127	5	Yes	Yes

Site Type Key: U: Urban, US: Urban and Suburban, S: Sub-Urban, RML: Rural Multilane, R2L: Rural 2-Lane 2-Way, 3SG: 3-Leg Signalized, 4SG: 4 Leg Signalized, 3ST: 3-Leg Minor Stop-Controlled, 4ST: 4-Leg Minor Stop Controlled, 4AWST: 4-Leg All-Way Stop

In contrast, a Florida study showed the base HSM models to underestimate fatal and injury crashes by a factor of two (Srinivasan et al. 2011) while SPFs that were calibrated for intersections in Ohio showed significant under-prediction at urban three-leg and four-leg signalized intersections (Troyer et al. 2015). Research in Saskatchewan (Young and Park 2012) showed the HSM SPFs to typically under-predict crashes across the three intersection types examined. Additional international work in Brazil explored the transferability of HSM models to urban intersections (Cunto et al. 2015). The results suggest that the calibrated HSM baseline SPFs should be used with caution, with the authors noting the importance of analyzing the effects of the calibration sample size on model stability. Ultimately, it has been postulated that the differences in calibration factors are reflective of differences between individual jurisdictions and those states where the HSM models were developed (Shin et al. 2014; Sun et al. 2014).

Given the significant variability in predictive performance across regions, a number of states have developed SPFs specific to their jurisdictions. Virginia is one of several states that have conducted extensive research on SPFs, including the development of SPFs for 3-leg and 4-leg signalized and stop-controlled intersections in urban and rural areas. Separate SPFs were developed on statewide basis, as well as at the regional-specific (Northern, Western, and Eastern regions) level to account for differences in various geographic areas of the state (Garber and Rivera 2010).

Research in Colorado resulted in the development of SPFs for ten types of urban intersections, including separate SPFs for total and injury crashes (Persaud and Lyon 2009). SPFs were developed in Oregon for eight intersection types based on traffic control, land use, and number of legs (Monsere et al. 2011). These categories were chosen to align with the intersection types in the HSM.

A recent study in Pennsylvania (Donnell et al. 2014) examined rural two-lane intersections. SPFs were developed for three-leg and four-leg intersections with both signal and minor street stop-control. SPFs were also estimated for four-leg all-way stop controlled intersections on two-lane rural roads.

Collectively, the domestic and foreign studies have indicated that direct application of the SPFs from the HSM (or other non-local source) does not tend to provide accurate results without either careful calibration or re-estimation using local data. Consequently, the primary purpose of this study was to develop a series of SPFs and other safety tools that can be used by the Michigan Department of Transportation (MDOT) as a part of their continuing traffic safety efforts.

2.2.2 Areas of Research Need

One area in SPF application where minimal research has been conducted is the utilization of geographic indicator variables as crash modification factors in the SPF model. This methodology has been utilized in several studies (Savolainen et al. 2015, Kweon et al. 2014) where binary indicator variables were used to create CMFs for each of the geographic administrative regions of the state department of transportation. These geographic indicators are then be utilized to account for variations in geometric design standards, driver behaviors, and crash reporting differences between specific geographic locations. This research utilized a vast array of explanatory variables in conjunction with the regional indicator variables to examine if the inclusion of regional indicators inhibits the ability to identify geometric variables which impact traffic crashes.

Second, this study investigates the use of random effects models (sometimes referred to as random intercept or mixed models) to develop safety performance functions. Specifically of interest is whether the same explanatory variables are identified by modeling traffic crashes on an aggregated data set versus on a disaggregated version.

2.3 Data Description and Data Collection Process

Ultimately, the accuracy of an SPF depends largely on the quality of the data from which it is developed. The development of robust SPFs requires a crash database that is comprehensive and includes information on specific crash location, collision type, severity, relationship to junction, and types of maneuvers of the involved vehicles. Roadway data is also important, including the physical features within the right-of-way. Roadway geometry data that are recommended for use in safety analyses include: lane width; shoulder width and type; horizontal curve length, radius, and superelevation; grade; driveway density; and indicator variables for features such as auxiliary turn lanes (AASHTO 2010).

In 2008, the Model Minimum Uniform Crash Criteria (MMUCC) guidelines were developed with funding provided by the National Highway Traffic Safety Administration (NHTSA) in collaboration with the Governor's Highway Safety Association (GHSA), Federal Highway Administration (FHWA), Federal Motor Carrier Safety Administration (FMCSA), State DOTs, law enforcement agencies, and other traffic safety stakeholders. The MMUCC consists of a recommended minimum set of data elements for States to include in their crash forms and databases (NHTSA 2008). This set includes 110 data elements, 77 of which are to be collected at the scene, 10 data elements to be derived from the collected data, and 23 data elements to be obtained after linkage to driver history, injury and roadway inventory data.

As a part of this study, a comprehensive checklist of important data elements to be collected for the purposes of SPF development was created. As a starting point, an inventory file was obtained from MDOT. This file included location information for the following four types of site locations:

- 3-leg signalized intersections
- 4-leg signalized intersections
- 3-leg intersections with stop-control on the minor approach
- 4-leg intersections with stop-control on the minor approaches

For the purposes of SPF development, the HSM suggests a minimum sample size of 30 to 50 sites, which collectively experience a minimum of 100 total crashes per year. For the purposes of this study, another objective was to provide SPFs that are able to account for important differences across each of MDOT's seven geographic regions. Consequently, the research began with the random selection of 50 intersections from each region within the four site types illustrated in Figure 1. This figure also indicates the total number of intersections maintained by MDOT according to this inventory file.

While 50 sites were identified within most regions and site types, there are several regions where sufficient numbers of sites were not available as shown in Table 2. This was particularly true for three-leg signalized intersections as there are only 485 such locations across Michigan.

Table 2. Sites by MDOT Region and Intersection Type

Intersection	MDOT Region							
Type	Superior	North	Grand	Bay	Southwest	University	Metro	Total
3SG	9	24	26	21	38	38	55	211
3ST	50	51	51	50	51	50	50	353
4SG	48	50	51	50	52	50	50	351
4ST	50	50	50	50	50	50	50	350

3-Leg, Minor Approach Stop-Controlled
(N = 5,731 Sites)

3-Leg Signalized
(N = 485 Sites)

4-Leg, Minor Approach Stop-Controlled
(N = 2,695 Sites)

4-Leg Signalized
(N = 1,710 Sites)

Figure 1. Intersection Site Types

Once intersections were identified within each of the seven regions and four site types, data were collected from existing data sources that were either available publicly or through MDOT. These data sources included the following databases and files:

- Michigan State Police Statewide Crash Database;
- MDOT SafetyAnalyst Calibration File;
- Michigan Geographic Data Library (MiGDL) All Roads File;
- MDOT SafetyAnalyst Annual Average Daily Traffic File; and
- MDOT Sufficiency File.

A quality assurance/quality control (QA/QC) process was implemented to verify the data in these sources using MDOT's online linear referencing tool, *PR Finder,* which allows users to identify locations based on the Physical Road (hence the PR) and mile point. Detailed aerial photography was then reviewed using Google Earth. Further details of each respective data source is provided in the following sections of this document.

2.3.1 Michigan State Police Statewide Crash Database

The Michigan State Police (MSP) crash database contains details of all reported crash records in the state of Michigan. Records in this database are maintained at the crash-, vehicle-, and person-levels. For the purposes of this research, 14 crash level fields from the database were obtained. The fields that were collected are defined here:

- crsh_id- unique identifier for each crash, and was used as the basis for linking the spreadsheets
- date_val-contains the date the crash occurred, which allowed the crash to be assigned to a particular year
- fatl_crsh_ind-identifies the crash as having at least one fatality
- num_injy_a-total number of people sustaining "A level" injuries in the crash
- num_injy_b-total number of people sustaining "B level" injuries in the crash
- num_injy_c-total number of people sustaining "C level" injuries in the crash
- prop_damg_crsh_ind-identifes the crash as being property damage only (PDO)
- crsh_typ_cd-defines the crash as single-vehicle or one of nine multiple-vehicle collision types
- rdwy_area_cd-indicates where on the roadway a crash occurred, only crashes with codes relavent to intersections were considered
- ped_invl_ind-indicates that a pedestrian was involved in the crash
- bcyl_invl_ind-indicates that a bicycle was involved in the crash
- intr_id-assigns the crash to a specific intersection node in the Calibration file
- crnt_x_cord-the longitude at which the crash occurred
- crnt_y_cord-the latitude at which the crash occurred

As was previously mentioned, this crash was focused on "crash" level data. Crashes were defined based on the most significant injury sustained by anyone involved in the crash. Crashes involving bicycles or pedestrians were separated from vehicle-only crashes for the purpose of the data analysis.

2.3.2 Michigan DOT Calibration File

A comprehensive database of potential intersections to be considered in this study, was furnished by MDOT. The file contained four spreadsheets relevant to this study, titled 3ST, 3SG, 4ST, and 4SG. In all, 12,241 locations were identified in the file. In addition to identifying the sites, the file contained some information for each location, the most useful of which is described below:

- A unique identifier corresponding to the crash database
- PR|PRMP-the location of the intersection based on Michigan's linear referencing scheme "Physical Road" and "Physical Road Mile Point"

The information contained in this file was ultimately used as the basis for the selection of locations included in this study. Although additional information in the file is potentially useful, some problems arise when trying to use it. First, directions are of little concern for the creation of SPFs, while information such as which is the major leg and which is the minor leg is much more useful. Second, each entry in the file does not necessarily correspond to a complete intersection, but just a node in a link node network. The intersection of a boulevard with a two-way street is typically represented by two nodes, meaning that many of the entries in the file must be paired with another entry. Much of the information in this file was captured in more detail during a thorough data collection process leading to the creation of the final data set.

2.3.3 Michigan Geographic Data Library (MiGDL) All Roads file

In order to facilitate the use of GIS software for this project, a GIS shapefile was obtained from the Michigan Geographic Data Library from the Michigan Center for Geographic

Information (MCGI) website. The file consists of all the road segments found statewide. Although the file has a total of 36 attribute fields, the following three were of particular use for this project:

- PR-Physical Road ID number
- BMP-Beginning PR mile point for linear referencing system
- EMP-Ending PR segment mile point

2.3.4 Annual average daily traffic estimates from MDOT Safety Analyst file

A zip file containing traffic volume used by MDOT's SafetyAnalyst was used as the source for AADT information for this project. A .csv file was extracted from the zip file containing major and minor road AADT information for 34,915 nodes for the years 2000-2012. In addition to this information, the file also contained several identification fields identified below.

- A unique identifier
- A designation of US, state, interstate, etc. highway and route number
- Identification of the SafetyAnalyst subtype
- The names of the two intersecting roads
- The PR and PRMP of the intersection on the major and minor roads
- The cardinal direction of the roadway

2.3.5 MDOT Sufficiency File

MDOT sufficiency files were made available for the years 2004 through 2012. The sufficiency files contain 122 fields for the state maintained roads in Michigan. The data is broken into segments of varying length. As the research ultimately involved a detailed site review of each intersection the Sufficiency file was primarily used to determine major road speed limits as a part of this study.

2.3.6 Construction of the Preliminary Dataset

For the purposes of this analysis, a study period from 2008 to 2012 was considered, based on the availability of data at the beginning of the project. To assemble the data set, the observations in the intersection specific tabs of the calibration file were converted into one large list of 12,241 locations. AADT was then linked to each intersection using the PR and PRMP fields.

Crashes were queried from the MSP crash database for each of the 12,241 nodes in the MDOT Calibration file by matching the unique identifier fields.

This crash query was exported as an Excel file containing the 14 fields discussed in the MSP Crash Database section. A threshold value of 0.04 miles was established as the maximum distance from an intersection node that a crash would be considered an "intersection" crash, requiring the mapping of the intersections and crashes using GIS software. The All Roads file was used as the framework for the map. Linear referencing was utilized to locate the intersection nodes (which did not have coordinates for location) on Michigan's roadway network by Physical Road (PR) and Mile Point (MP). Crashes were then added to the map by latitude and longitude coordinates included in the crash report. To exclude crashes that were outside of the established 0.04 miles, a buffer was used around each of the intersection nodes.

The crashes that were within 0.04 miles of an intersection node were then tabulated by year, type, and severity so that each node would have a count of the crashes that occurred near it by type and severity. Of the 12,241 nodes provided in the Calibration file, 12,170 were able to be paired with AADT from the Safety Analyst file and mapped onto the All Roads file, with 71 nodes having a PR or PRMP that did not correspond to either the Safety Analyst AADT file or the All Roads file.

While many of the aforementioned intersection nodes were representative of a complete intersection, many others were a portion of a more complex intersection such as a boulevard intersecting a two-way street, or the intersection of two boulevards, as shown in Figure 2.

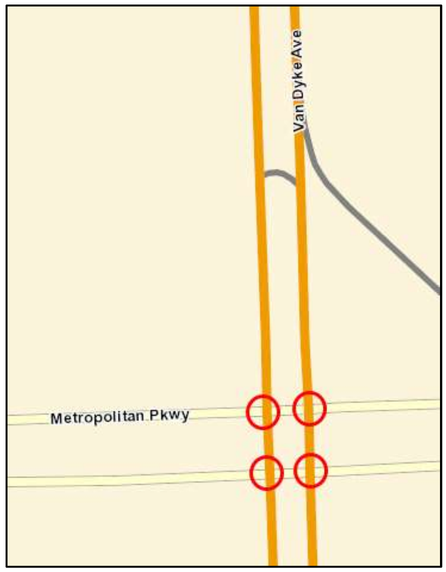


Figure 2. Boulevard-Style Four-Node Intersection

Other nodes were not intersections at all, but the beginning or end point of a boulevard, the location of a median turnaround (Michigan Left), or the location of yield-controlled or uncontrolled merging and diverging lanes as shown in Figure 3.

Figure 3. Example of Merge/Diverge Point Classified as an Intersection

This necessitated an exhaustive QA/QC of the data to join nodes of the same intersection together, as well as to remove the nodes that would not be considered an intersection from the dataset. Utilizing a logic function in Microsoft Excel, the names of the intersecting roadways from the AADT file were used so that nodes potentially belonging to the same intersection were identified. The PR Finder was used to locate the sites and view initial satellite imagery, with Google Earth providing additional satellite imagery. Images were reviewed to verify whether nodes were properly identified as a complete intersection. Nodes that were found not to be an intersection were excluded from further analysis, leaving a final data set consisting 10,621 intersections. In order for the properly linked intersection nodes to have characteristics representative of the entire intersection, the crashes assigned to these nodes were summed, as was the AADT for each side of a boulevard; non-boulevard streets had their AADT values averaged. Table 3 provides details of the resulting data set, including a count of the number of intersections by type, as well as averages of the major AADT, minor AADT, and total annual crashes.

Table 3. Average Major AADT, Minor AADT and Annual Crashes by Intersection Type

	3SG	3ST	4SG	4ST
Number of Intersections	485	5,731	1,710	2,695
Average Major Road AADT	20,709	15,985	23,892	14,571
Average Minor Road AADT	4,967	1,234	9,547	1,776
Average Annual Crashes	2.67	0.42	7.78	1.05

2.3.7 Manual Data Collection and Review

In order to create a data set containing geometric data (e.g. road width, number of lanes), as well as road use characteristics (e.g. bus stops, roadside parking), a detailed site review was conducted utilizing Google Earth and the MDOT PR Finder. Detailed geometry and site characteristics data were obtained, with the following list summarizing the data collection process:

- Number of Lanes: The number of lanes was determined for each approach and receiving leg. This information was disaggregated into the number of exclusive left-turn lanes, exclusive right-turn lanes, and through lanes. While both the entry approach and receiving lanes were reviewed, only the inbound lanes were considered for the purpose of the subsequent analysis.
- Road widths: the widths of intersecting roads were measured from curb to curb for all approaches. For the purpose of analysis, if both legs were present these measured widths were averaged along the same street, otherwise the measured values were directly used.
- Skew angle: The skew angle for each intersection was calculated as the smallest absolute difference between the headings of the intersecting approaches. The smallest angle was the variable of interest since it is the controlling situation where the available sight distance is minimum resulting in greater potential for crash occurrence. A sample skew angle measurement is shown in Figure 4.

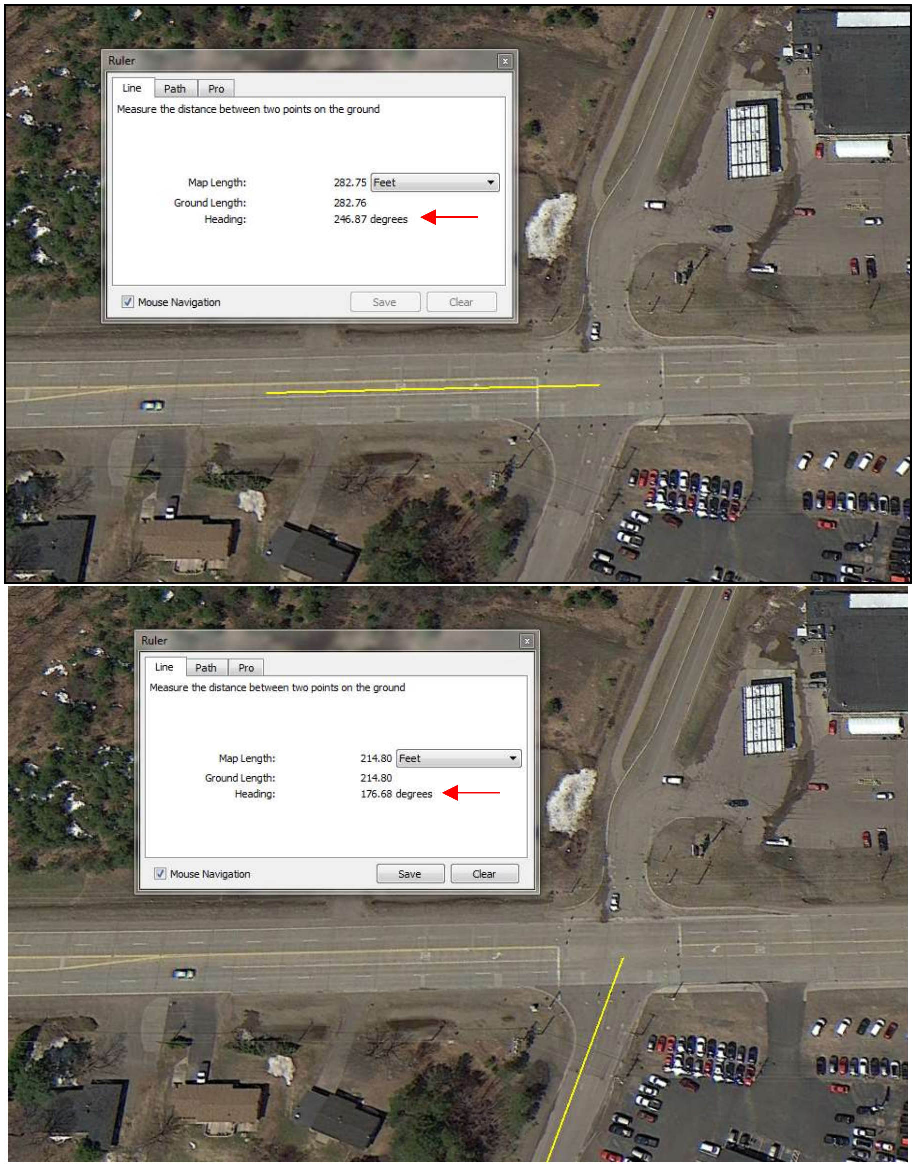


Figure 4. Skew Angle Measurement Example

- Number of driveways: The total number of driveways were collected on both sides of the intersecting streets up to a distance of 0.04 miles from the center of the intersection along both the major and minor street.
- Bike Lanes and Roadside Parking: Presence of exclusive bike lanes and roadside parking was also specified.
- Bus Stops: Presence of bus stops within a distance of 1000 feet from the center of the intersection was investigated both for the major and minor road. Although bus stops are usually depicted on Google Maps, not all the bus stops can be located using the aerial view. Hence, more detailed exploration through Google Street View was required.

- Schools: A distance of 0.5 mile from the center of the intersection was used to explore for schools both on major and minor road. As with bus stops, Street View was used for verification where the aerial view was unclear. This field includes K-12 schools, as well as universities and colleges.
- Pedestrian features: The presence of sidewalks and ADA ramps was specified both for the major and minor road.
- Median Turn-around (MTA): A median turn-around refers to the case where, near an intersection, at least one road is a divided boulevard and left-turns onto the divided highway are prohibited. In such instances, left-turns are generally accommodated by vehicles making a right-turn, followed by a U-turn through the median as shown in Figure 5. All such instances were indicated for vehicles attempting to turn left from both the major and minor road.
- Distance of MTA: In cases were the presence of a median turn around was specified, its distance from the center of the intersection was also measured.

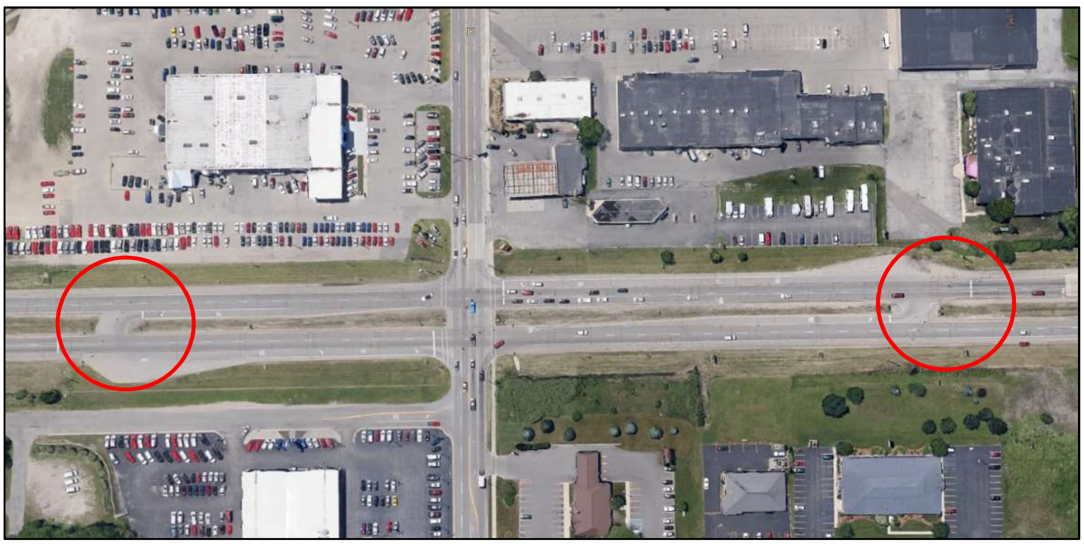


Figure 5. Median Turn-around Field Example

- Presence and length of storage lanes: The presence and lengths of storage lanes were determined as illustrated in Figure 6. For the case of intersections with two-way left-turn lanes (TWLTLs), no storage length was specified, though the presence of the turn lane was indicated in the database.

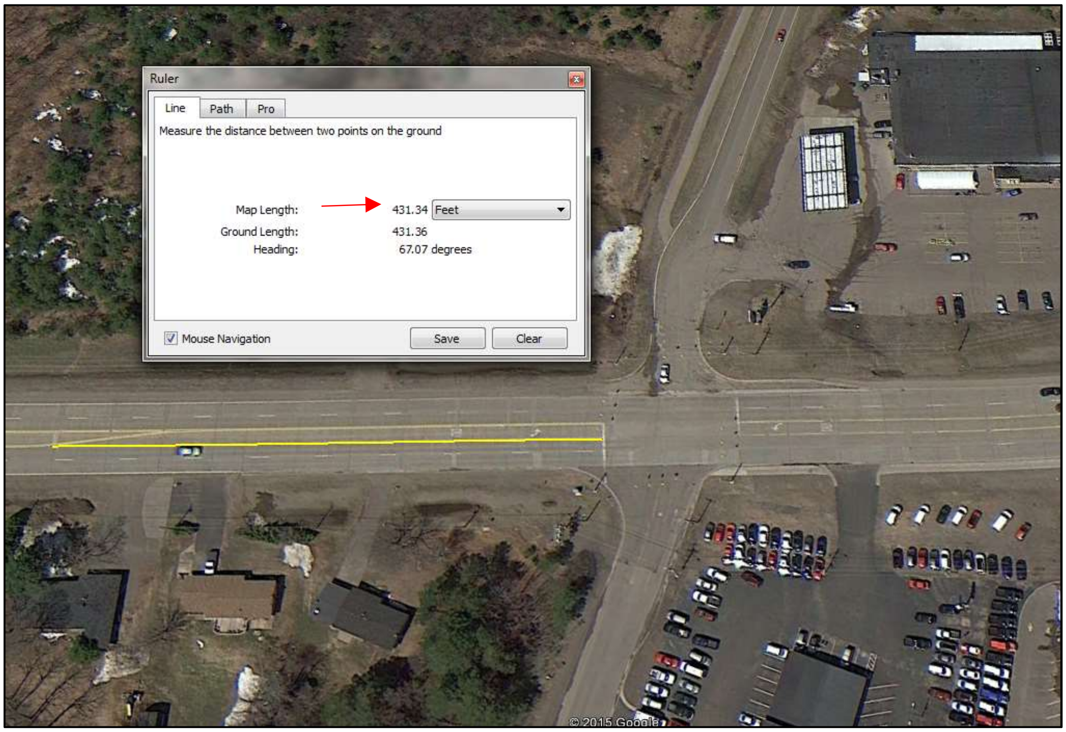


Figure 6. Storage Lane Length Measurement Example

- Median Types: In cases where either of the major or minor road was divided the type of the median was also identified. Medians were classified into different categories including curbed, curbed with grass, curbed with grass and vegetation, grass only, concrete barrier, guardrail barrier, and asphalt medians.
- Right-Turn-on-Red (RTOR): This field indicates those signalized intersections where vehicles are allowed to turn right while the signal head is red.
- Flashing Beacon: Those intersections where a flashing beacon is installed as well as/instead of a stop sign were flagged during the data collection process.

Table 4 and Table 5 provide summary statistics for all relevant variables among the stop-controlled and signalized intersection databases, respectively. Each table presents the minimum, maximum, and mean values, along with the standard deviation for each variable. Traffic crashes are more frequent at signalized intersections as opposed to unsignalized crashes, as well as at four-leg intersections as opposed to three-leg. This is due to the volume of traffic at each facility type, which follows the same general trend.

Table 4: Descriptive Statistics for variables of interest for Stop-Controlled Intersections

Intersection Type	3-Leg Minor Stop Controlled				4-Leg Minor Stop Controlled			
Variable	Min	Max	Mean	Std. Dev.	Min	Max	Mean	Std. Dev.
Maj Rd AADT	97.0	48824.0	13040.9	7541.93	929.00	50206.0	13618.02	7913.7
Min Rd AADT	42.5	11630.5	516.65	965.36	85.00	44209.0	1898.66	3409.9
Maj Rd Thru Lanes	2.00	8.00	3.07	1.07	1.00	6.00	3.12	1.09
Maj Left Turn Lanes	0.00	2.00	0.58	0.81	0.00	2.00	0.99	0.98
Min Rd Thru Lanes	0.00	2.00	0.96	0.24	0.00	4.00	1.98	0.26
Skew	0.01	69.33	7.89	12.13	0.00	64.00	8.41	12.63
Lighting Presence	0.00	1.00	0.72	0.45	0.00	1.00	0.76	0.43
Maj Driveway Count	0.00	15.00	2.87	2.49	0.00	10.00	2.19	2.18
Min Driveway Count	0.00	8.00	1.70	1.39	0.00	10.00	2.47	2.30
Maj Sidewalk Presence	0.00	1.00	0.59	0.49	0.00	1.00	0.71	0.46
Min Sidewalk Presence	0.00	1.00	0.40	0.49	0.00	1.00	0.67	0.47
Ped Ramp Presence	0.00	1.00	0.52	0.50	0.00	1.00	0.45	0.50
Maj Rd Width	24.76	155.24	52.77	19.08	22.00	171.00	51.26	18.58
Min Rd Width	12.60	115.84	30.39	10.64	14.00	65.50	30.79	7.37
Maj Rd Bike Lanes	0.00	1.00	0.00	0.07	0.00	1.00	0.03	0.16
Min Rd Bike Lanes	0.00	1.00	0.00	0.05	0.00	1.00	0.01	0.08
Maj Rd Bus Stop	0.00	1.00	0.18	0.38	0.00	1.00	0.20	0.40
Min Rd Bus Stop	0.00	1.00	0.04	0.19	0.00	1.00	0.01	0.12
Maj Rd Parking	0.00	1.00	0.07	0.25	0.00	1.00	0.16	0.36
Min Rd Parking	0.00	1.00	0.21	0.41	0.00	1.00	0.55	0.50
Maj Rd Median	0.00	1.00	0.02	0.13	0.00	1.00	0.04	0.20
Min Rd Median	0.00	1.00	0.06	0.24	0.00	1.00	0.01	0.11
Within 1/2 mile of K-12 school	0.00	1.00	0.34	0.47	0.00	1.00	0.27	0.44
Within 1 mile of non-motorized path	0.00	1.00	0.30	0.46	0.00	1.00	0.30	0.46
Superior Region	0.00	1.00	0.14	0.35	0.00	1.00	0.14	0.35
North Region	0.00	1.00	0.14	0.35	0.00	1.00	0.14	0.35
Grand Region	0.00	1.00	0.14	0.35	0.00	1.00	0.14	0.35
Bay Region	0.00	1.00	0.14	0.35	0.00	1.00	0.14	0.35
Southwest Region	0.00	1.00	0.14	0.35	0.00	1.00	0.14	0.35
University Region	0.00	1.00	0.14	0.35	0.00	1.00	0.14	0.35
Metro Region	0.00	1.00	0.14	0.35	0.00	1.00	0.14	0.35
Maj Rd Speed Limit	25.00	55.00	43.07	9.05	25.00	65.00	38.70	9.21
Maj Rd One-Way	0.00	1.00	0.15	0.36	0.00	1.00	0.12	0.33
Min Rd One-Way	0.00	0.00	0.00	0.00	0.00	1.00	0.02	0.15
Total Crashes	0.00	19.00	0.49	1.25	0.00	24.00	1.23	1.83
Fatal/Injury Crashes	0.00	6.00	0.10	0.40	0.00	7.00	0.32	0.67
PDO Crashes	0.00	18.00	0.39	1.08	0.00	21.00	0.91	1.49

Table 5: Descriptive Statistics for Variables of interest for Signalized Intersections

Intersection Type	3-Leg Signalized				4-Leg Signalized			
Variable	Min	Max	Mean	Std. Dev.	Min	Max	Mean	Std. Dev.
Maj Rd AADT	4391.0	62094.0	20012.1	10001.9	4033.00	120082.0	21159.0	15155.6
Min Rd AADT	45.00	42828.00	3810.40	4911.33	88.00	69321.00	8901.74	7999.21
Maj Rd Thru Lanes	0.00	10.00	3.60	1.29	0.00	10.00	3.64	1.24
Maj Left Turn Lanes	0.00	2.00	1.06	0.77	0.00	4.00	1.35	0.93
Min Rd Thru Lanes	0.00	4.00	0.46	0.76	0.00	8.00	2.64	1.07
Skew	0.00	74.34	10.21	14.51	0.00	61.04	9.85	14.25
Lighting Presence	0.00	1.00	0.73	0.44	0.00	1.00	0.96	0.20
RTOR Permitted	0.00	1.00	0.90	0.31	0.00	1.00	0.91	0.29
Maj Driveway Count	0.00	10.00	2.38	2.18	0.00	13.00	3.40	2.71
Min Driveway Count	0.00	7.00	1.34	1.41	0.00	14.00	3.74	2.61
Maj Sidewalk Presence	0.00	1.00	0.71	0.45	0.00	1.00	0.79	0.40
Min Sidewalk Presence	0.00	1.00	0.63	0.48	0.00	1.00	0.77	0.42
Ped Ramp Presence	0.00	1.00	0.67	0.47	0.00	1.00	0.83	0.37
Maj Rd Width	25.42	282.62	69.21	35.27	27.91	314.48	65.74	29.50
Min Rd Width	14.77	176.45	45.17	16.58	25.99	188.98	50.23	21.68
Maj Rd Bike Lanes	0.00	1.00	0.01	0.12	0.00	1.00	0.03	0.16
Min Rd Bike Lanes	0.00	1.00	0.01	0.10	0.00	1.00	0.02	0.14
Maj Rd Bus Stop	0.00	1.00	0.26	0.44	0.00	1.00	0.31	0.46
Min Rd Bus Stop	0.00	1.00	0.06	0.23	0.00	1.00	0.20	0.40
Maj Rd Parking	0.00	1.00	0.09	0.28	0.00	1.00	0.14	0.35
Min Rd Parking	0.00	1.00	0.15	0.36	0.00	1.00	0.15	0.36
Maj Rd Median	0.00	1.00	0.12	0.32	0.00	1.00	0.11	0.31
Min Rd Median	0.00	1.00	0.10	0.29	0.00	1.00	0.04	0.20
Within 1/2 mile of K-12 school	0.00	1.00	0.16	0.37	0.00	1.00	0.40	0.49
Within 1 mile of non-motorized path	0.00	1.00	0.32	0.47	0.00	1.00	0.33	0.47
Superior Region	0.00	1.00	0.04	0.20	0.00	1.00	0.13	0.34
North Region	0.00	1.00	0.11	0.32	0.00	1.00	0.14	0.35
Grand Region	0.00	1.00	0.12	0.33	0.00	1.00	0.15	0.35
Bay Region	0.00	1.00	0.10	0.30	0.00	1.00	0.14	0.35
Southwest Region	0.00	1.00	0.18	0.38	0.00	1.00	0.15	0.36
University Region	0.00	1.00	0.18	0.38	0.00	1.00	0.14	0.35
Metro Region	0.00	1.00	0.26	0.44	0.00	1.00	0.14	0.35
Maj Rd Speed Limit	25.00	55.00	41.05	8.41	25.00	70.00	38.70	9.44
Maj Rd One-Way	0.00	1.00	0.08	0.27	0.00	1.00	0.12	0.32
Min Rd One-Way	0.00	1.00	0.09	0.29	0.00	1.00	0.11	0.31
Total Crashes	0.00	30.00	3.90	3.82	0.00	61.00	8.41	8.31
Fatal/Injury Crashes	0.00	6.00	0.88	1.14	0.00	13.00	1.93	2.13
PDO Crashes	0.00	26.00	3.02	3.19	0.00	51.00	6.48	6.82

2.4 Study Methods

2.4.1 Visual Data Analysis

After the database was assembled, a series of preliminary analyses were conducted to examine general trends across the sample of study locations. This included assessing the univariate relationships between traffic crashes and each prospective predictor variable. Correlation among predictor variables was also examined and helped to inform the subsequent estimation of the SPFs. To this end, visual analysis of the data was employed. Figure 7 shows the relationship between the number of crashes (all severities) and the annual average daily traffic (AADT) for the major approaches for signalized intersections, while Figure 8 conveys the same information for unsignalized intersections. These figures show that a non-linear relationship generally exists between traffic flow and the number of crashes. Crashes are shown to increase less rapidly at higher volumes, which is consistent with prior research in this area.

When examining the figure, there are several intersection locations that experienced significantly higher or lower numbers of crashes over the study period. As a part of the data collection process, careful quality assurance and quality control procedures were followed. This included a review of these potential outliers. Ultimately, all of the intersections included in the study were similar in terms of their geometric and traffic control characteristics. No sites were removed on the basis of their crash history during the study period. It is important to note that these figures represent only the effects of major road traffic volumes. Consequently, the effects of other important predictor variables are not reflected here. As an example, fewer crashes tended to be observed at locations with medians or where specific turning movements were prohibited. This explains several of the high volume locations that experienced fewer crashes on average.

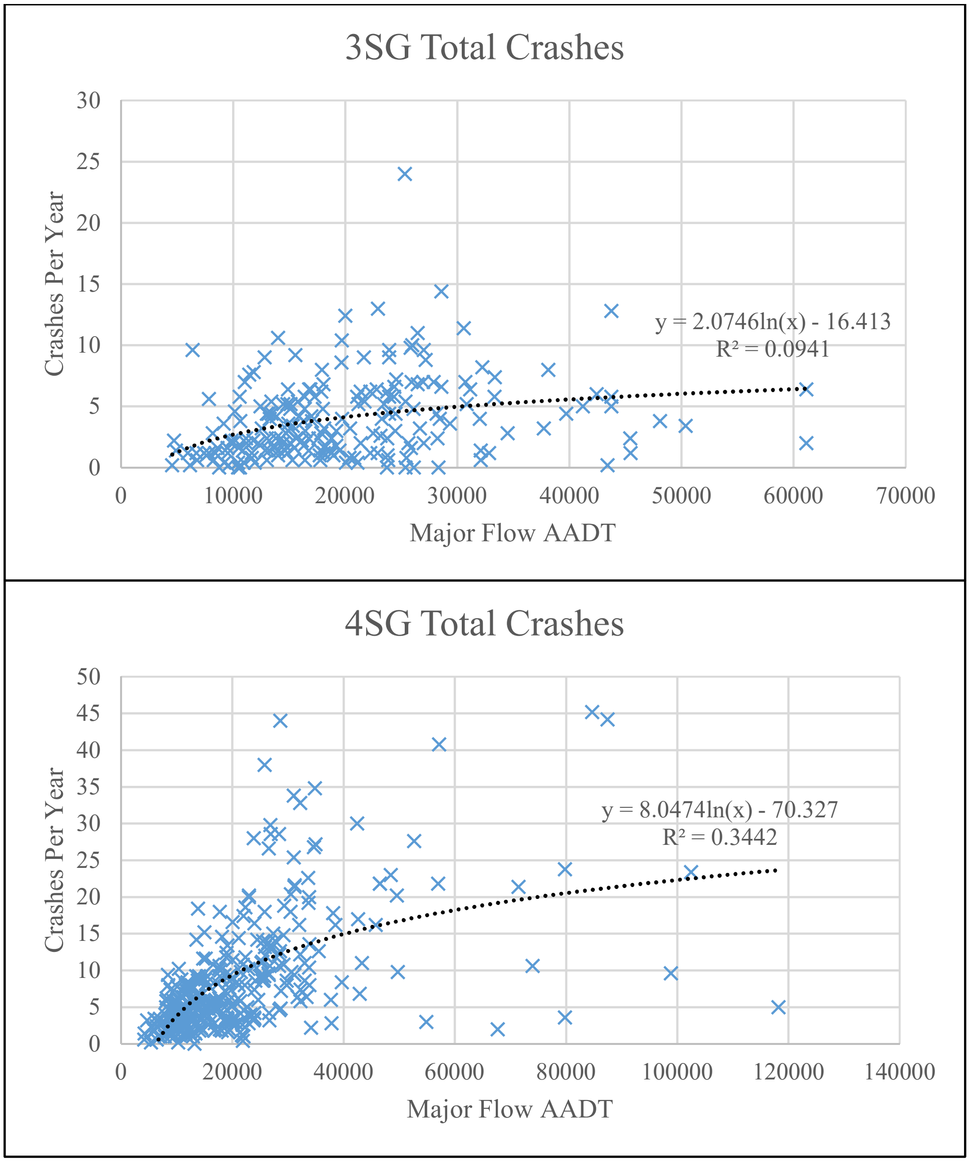


Figure 7. Relationship Between the Number of Vehicle-Only Crashes and Major flow AADT for Signalized Intersections.

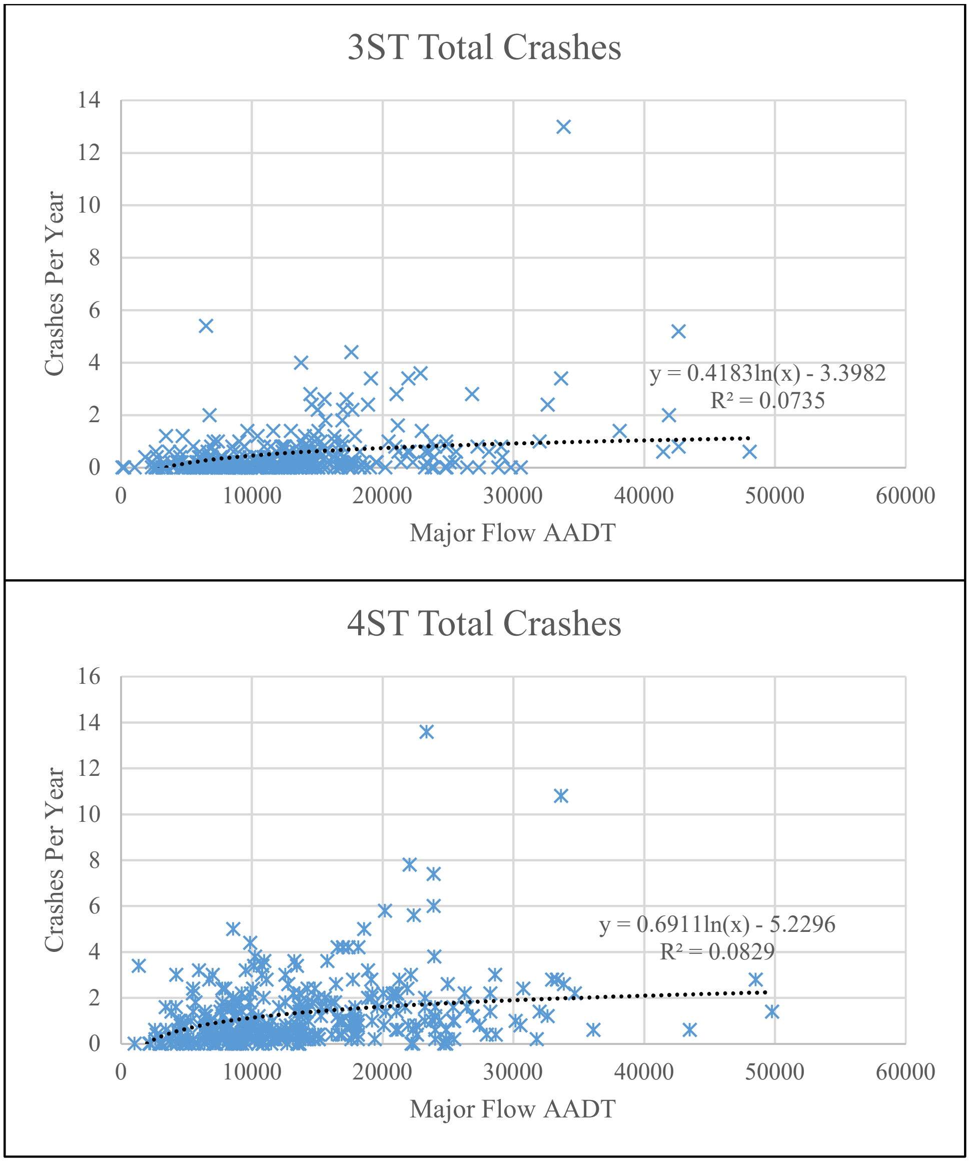


Figure 8. Relationship Between the Number of Vehicle-Only Crashes and Major flow AADT for Unsignalized Intersections

2.4.2 Development of Safety Performance Functions

After examining these general relationships between crashes and traffic volumes within each of the four site types, a series of SPFs were developed at varying degrees of complexity. These SPFs take the form of generalized linear models. As crash data are comprised of non-negative integers, traditional regression techniques (e.g., ordinary least-squares) are generally not appropriate. Given the nature of such data, the Poisson distribution has been shown to provide a better fit and has been used widely to model crash frequency data. In the Poisson model, the probability of intersection i experiencing y_i crashes during a one-year period is given by Equation 4,

$$P(y_i) = \frac{EXP(-\lambda_i)\lambda_i^{y_i}}{y_i!} \tag{4}$$

where P(yi) is probability of intersection i experiencing yi crashes and λ_i is the Poisson parameter for intersection i, which is equal to the segments expected number of crashes per year, E[yi]. Poisson models are estimated by specifying the Poisson parameter λ_i (the expected number of crashes per period) as a function of explanatory variables, the most common functional form being given by Equation 5,

$$\lambda_i = exp(\beta X_i) \tag{5}$$

where X_i is a vector of explanatory variables and β is a vector of estimable parameters.

A limitation of this model is the underlying assumption of the Poisson distribution that the variance is equal to the mean. As such, the model cannot handle overdispersion wherein the variance is greater than the mean. Overdispersion is common in crash data and may be caused by data clustering, unaccounted temporal correlation, model misspecification, or ultimately by the nature of the crash data, which are the product of Bernoulli trials with unequal probability of events (Lord 2006). Overdispersion is generally accommodated through the use of negative binomial models (also referred to as Poisson-gamma models).

The negative binomial model is derived by rewriting the Poisson parameter for each intersection as shown in Equation 6,

$$\lambda_i = exp(\beta X_i + \varepsilon_i) \tag{6}$$

where $EXP(\varepsilon_i)$ is a gamma-distributed error term with mean 1 and variance α. The addition of this term allows the variance to differ from the mean as shown in Equation 7:

$$VAR[y_i] = E[y_i] + \alpha E[y_i]^2 \tag{7}$$

The negative binomial model is preferred over the Poisson model since the latter cannot handle overdispersion and, as such, may lead to biased parameter estimates (Lord and Park 2008). Consequently, the *HSM* recommends using the negative binomial model for the development of SPFs.

If the overdispersion parameter (α) is equal to zero, the negative binomial reduces to the Poisson model. Estimation of λ_i can be conducted through standard maximum likelihood procedures. While alternatives, such as the Conway-Maxwell model, have the advantage of accommodating both overdispersion and underdispersion (where the variance is less than the mean) (Lord and Mannering 2010), the negative binomial model remains the standard in SPF development.

Due to the presence of repeated observations resulting in temporal correlation among observations, random-effects were used to estimate models for the disaggregate data sets. The random effects framework allows the constant term β_0 of Equation 6 to vary as shown in Equation 8.

$$\beta_{0i} = \beta_0 + \omega_i \tag{8}$$

where the subscript i indexes a specific intersection $\omega_i$is normally distributed with mean zero and a variance that is estimated as a model parameter that is allowed to vary across intersections. By allowing the constant term to vary, unobserved heterogeneity unique to each location resulting from repeated observations. The constant term can therefore take on a higher or lower value at locations where crash frequency may be affected by the presence of variables not available in the data set.

By allowing the intercept term to vary from site to site, the predictive ability of the model is improved as the site-specific intercept term obtained from the model effectively depends on the crash history at a given site. The predictive capability of the random effects framework is therefore not necessarily comparable to the predictive ability of the standard negative binomial framework alone. In this vein, a more applicable comparison of predictive ability can be made through the application of empirical Bayes (EB) methodology to the predicted crashes obtained from the standard negative binomial model, as described in the HSM. The EB method utilizes the overdispersion parameter, α to determine the weighed adjustment factor, w, which is then used to estimate the expected number of crashes at a given location when combining observed crash data with the number of crashes predicted by an SPF. The formula for this weighting factor is given in Equation 9:

$$w = \frac{1}{1+(\alpha \times N_{spf})} \tag{9}$$

where:

α = overdispersion parameter, and
N_{spf} = predicted number of crashes by SPF.

Upon determining w, the expected number of crashes can then be determined using Equation 10:

$$N_{expected} = w \times N_{spf} + (1 - w) \times N_{observed} \tag{10}$$

where:

$N_{expected}$ = expected number of crashes determined by the EB method,
w= weighted adjustment factor, and
$N_{observed}$ = observed number of crashes at a site.

For further details of the EB method, the reader is referred to the HSM (AASHTO 2010). Model goodness-of-fit metrics, such as log-likelihood and AIC, do not account for the application of EB to the standard negative binomial model. Therefore, in order to compare the negative binomial with EB approach to the random effects approach, the models can be assessed in terms of predictive ability. This study focuses on two metrics of model fit, Mean Absolute Deviance (MAD) and Mean Squared Predictive Error (MSPE) (Oh et al. 2003), which are shown in Equations 11 and 12, respectively:

$$MAD = \frac{1}{n}\Sigma_{i=1}^{n}|\hat{y}_i - y_i| \tag{11}$$

and

$$MSPE = \frac{1}{n}\Sigma_{i=1}^{n}(\hat{y}_i - y_i)^2 \tag{12}$$

where:

n = the number of observations,
i = the ith observation,
$\hat{y}_i$ =the predicted value of the ith observation, and
y = the observed value of the ith observation.

As noted previously, several SPFs were developed at a part of this research at varying degrees of complexity. The complexity of the SPFs is reflective, in part, on the underlying data requirements. MDOT collects or estimates AADT on its entire trunkline system on a regular basis. While this is not necessarily the case with minor roads that are not state-maintained, AADT estimates have been developed for all such roads and were provided in the SafetyAnalyst AADT File that was used for the purposes of this study. Consequently, these simpler AADT-only models will provide a viable short-term tool for use in high-level safety planning activities.

As a part of this study, SPFs were examined at four levels of detail:

- Uncalibrated HSM – The intersection models from Chapter 12 of the HSM were applied directly using traffic volume data for the study sites.
- Calibrated HSM – The predicted number of crashes based upon the SPFs from the HSM were calibrated based upon the observed crashes at the study sites.
- Michigan-Specific Models with AADT and Regional Indicators – A series of Michigan-specific models were developed using only AADT for the major and minor roads. A simple statewide model was estimated, as well as a similar model that included a series of binary indicator variable for each MDOT region.
- Fully Specified Michigan-Specific Models – A series of detailed models were subsequently developed in consideration of AADT, regional indicator variables, and a diverse range of geometric variables.

2.5 Results and Discussion

2.5.1 Comparison of Uncalibrated and Calibrated HSM Models

The base SPFs from Chapter 12 of the HSM were first applied to nearly all-inclusive datasets for each of the four intersection types. These base models utilize only the AADT for the major and minor road as input values. This was done as the data for specific geometric information for each site was not available at the system-wide level in MDOT's SafetyAnalyst files (subsequent analysis beyond the initial calibration utilizes a sample of the data for which geometric information was manually collected). Separate estimates were obtained for total crashes, property damage only (PDO) crashes, and fatal/injury (F/I) crashes.

After applying these models, the resulting estimates for each study location were then compared to the observed values. The ratio of the total observed crashes to the estimated crashes (from the base SPFs) for the entire sample is used to estimate a calibration factor, as shown in the Equation 13:

$$C = \frac{N_{observed}}{N_{SPF}} \tag{13}$$

where C is the calibration factor Nobserved is the total number of crashes observed for a given intersection type, and NSPF is the total number of crashes predicted for a given facility type. The calibration factor can then be combined with the predicted crash value to obtain a more accurate estimated as shown in Equation 14:

$$N_{calibrated,i} = N_{spf,i} * C \tag{14}$$

where $N_{calibrated,i}$ is the calibrated prediction of crashes at a given location, $N_{spf,i}$ is the uncalibrated predicted number of crashes at a a given location as given by the SPF, and C is the calibration factor obtained from Equation 13.

The calibration factors provide a measure of how close the base SPFs from the HSM fit the Michigan data. The calibration factor for each of the three models (i.e., total, PDO, and F/I) and each of the four site types (3SG, 3ST, 4SG, and 4ST) are presented in Table 6.

Table 6: Calibration Factors for HSM Models

	Intersection Types	**3SG**	**3ST**	**4SG**	**4ST**
Single-Vehicle	Total	0.950	0.266	0.977	0.333
	PDO	0.825	0.232	0.648	0.311
	Fatal-Injury	1.338	0.353	2.002	0.512
Multi-Vehicle	Total	0.876	0.294	1.094	0.469
	PDO	1.100	0.340	1.331	0.563
	Fatal-Injury	0.561	0.171	0.750	0.301

By briefly scanning the calibration factors for the *HSM* models, it is evident that the accuracy of the base SPFs from the HSM vary widely by site type, crash type, and crash severity level, thus suggesting that the estimation of Michigan-specific SPFs is appropriate for the given data sets. Calibration factors with values greater than one indicate the SPF tends to under predict, such as was observed for 3SG and 4SG fatal and injury crashes. Calibration factors with values less than one indicate that the default SPFs tend to over predict, as was observed for all stop-controlled SPFs. The differences are reflective of several factors, including state-specific differences (e.g., driver characteristics, road design standards, weather, etc.), as well as the fact that only AADT was considered (and not geometric or road use characteristics).

2.5.2 Simple Michigan-Specific Safety Performance Functions

Having established that the base SPFs from the HSM do not generally provide consistent fit across intersection types, crash types, and crash severity levels, a series of Michigan-specific SPFs have been developed. These SPFs were (or will be) developed in two general forms:

- Michigan-Specific Models with AADT and Regional Indicators – A series of Michigan-specific models were developed using only AADT for the major and minor roads. A simple statewide model was estimated, as well as a similar model that included a series of binary indicator variables for each MDOT region.
- Fully Specified Michigan-Specific Models – A series of detailed models were subsequently developed in consideration of AADT, regional indicator variables, and a diverse range of geometric variables.

2.5.2.1 AADT-Only SPFs

This section presents the results of separate SPFs for fatal and injury (F/I) crashes and property damage only (PDO) crashes for three-leg signalized intersections. In order to demonstrate how SPFs can capture regional differences that would not otherwise be observable, an AADT-only SPF as well as a SPF with regional indicators have been developed. These models account for general differences in safety performance across the seven MDOT regions due to crash reporting characteristics, driver behavior, and weather conditions, among other things. Most notably, the traffic volumes observed in the Metro region are typically higher than those in the rest of the state, meaning drivers may be more used to operating in congestion as well as capturing the non-linear relationship between volume and crash frequency. For these models, parameter estimates are provided for AADT on the major and minor road. In model with regional indicators, the Metro region serves as the baseline and indicator variables are then used to adjust the estimates to each of the other regions.

Table 7, Figure 9, Figure 10, and Figure 11present the AADT only SPFs for four-leg signalized (4SG) intersections. These locations showed crashes to increase much more rapidly with respect to major road AADT as compared to minor road AADT.

Table 7. AADT Only SPF for Crashes at 4SG Intersections

	Total		Fatal and Injury Crashes		Property Damage Only Crashes	
Variable	**Value**	**Std. Dev**	**Value**	**Std. Dev**	**Value**	**Std. Dev**
Intercept	-7.859	0.295	-7.902	0.387	-8.436	0.315
Major AADT	0.792	0.032	0.709	0.043	0.801	0.035
Minor AADT	0.238	0.017	0.173	0.024	0.263	0.019
Inverse Dispersion Parameter	0.352	0.017	0.339	0.032	0.380	0.019

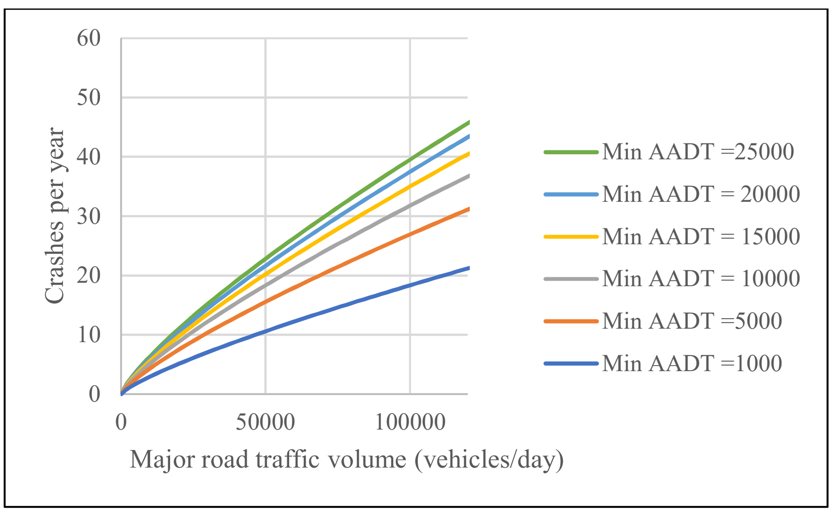


Figure 9. Graphical Form of Total Crash AADT Only SPF for Four-Leg Signalized (4SG) Intersections

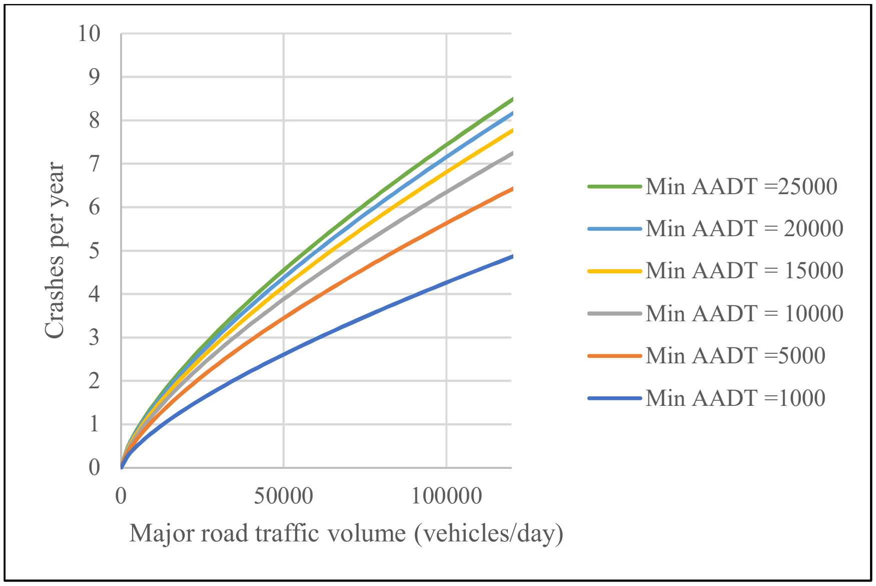


Figure 10. Graphical Form of Fatal and Injury Crash AADT Only SPF for Four-Leg Signalized (4SG) Intersections

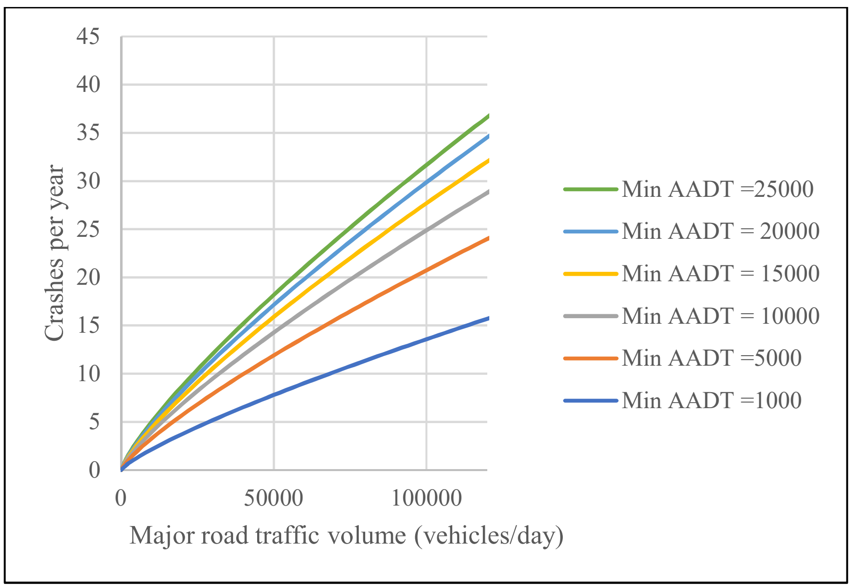


Figure 11. Graphical Form of PDO Crash AADT Only SPF for Four-Leg Signalized (4SG) Intersections

2.5.2.2 AADT and Regional Indicator SPFs

Table 8, Figure 12, Figure 13, and Figure 14 present the SPFs with regional indicators for three-leg signalized (4SG) intersections. Again, these locations showed crashes to increase much more rapidly with respect to major road AADT as compared to minor road AADT. When controlling for the effects of traffic volume, crashes were highest in the Superior and North regions and lowest in the Metro region. In each of the figures, the predicted crash value for each region is illustrated up to the maximum traffic volume observed in that region. The regional differences are therefore largely explained by differences in volume in the Metro region in comparison to the other regions. It is also possible that additional factors, such as localized driving behavior, weather, and crash reporting practices also play some role.

Table 8. SPF for Crashes at 4SG Intersections with AADT and Regional Indicators

	Total		Fatal and Injury Crashes		Property Damage Only Crashes	
Variable	**Value**	**Std. Dev**	**Value**	**Std. Dev**	**Value**	**Std. Dev**
Intercept	-10.045	0.365	-9.599	0.510	-10.696	0.392
Major AADT	0.953	0.036	0.841	0.050	0.967	0.039
Minor AADT	0.253	0.017	0.183	0.024	0.277	0.018
Added effect of Superior region	0.616	0.075	0.522	0.104	0.620	0.080
Added effect of North region	0.539	0.066	0.273	0.093	0.593	0.071
Added effect of Grand region	0.384	0.061	0.299	0.082	0.399	0.065
Added effect of Bay region	0.505	0.067	0.363	0.093	0.520	0.072
Added effect of Southwest region	0.749	0.064	0.393	0.089	0.827	0.068
Added effect of University region	0.451	0.065	0.309	0.090	0.472	0.070
Inverse Dispersion Parameter	0.313	0.016	0.327	0.031	0.329	0.018
*Note: Metro region serves as baseline reference category						

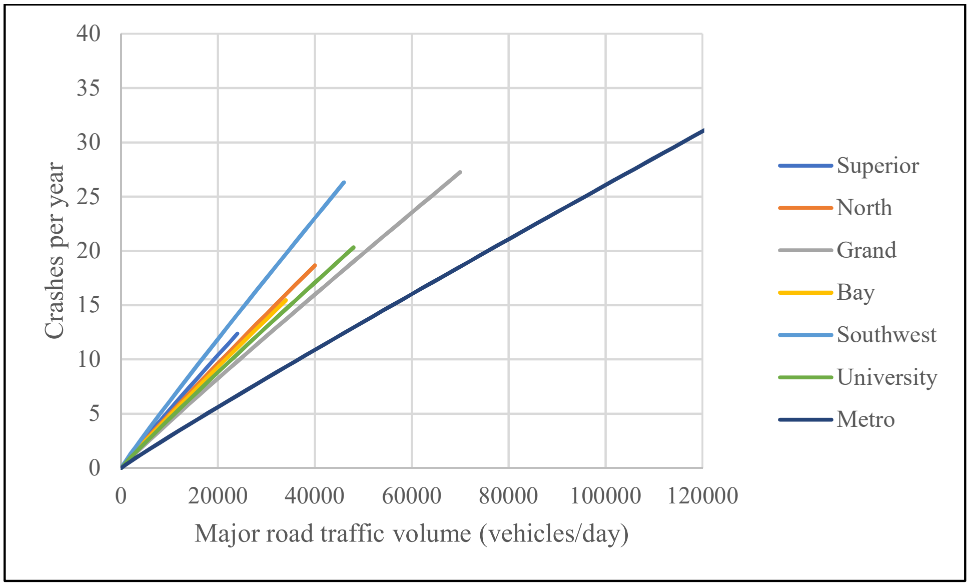


Figure 12. Graphical Form of Total Crash AADT and Regional Indicator SPF for Three-Leg Signalized (4SG) Intersections

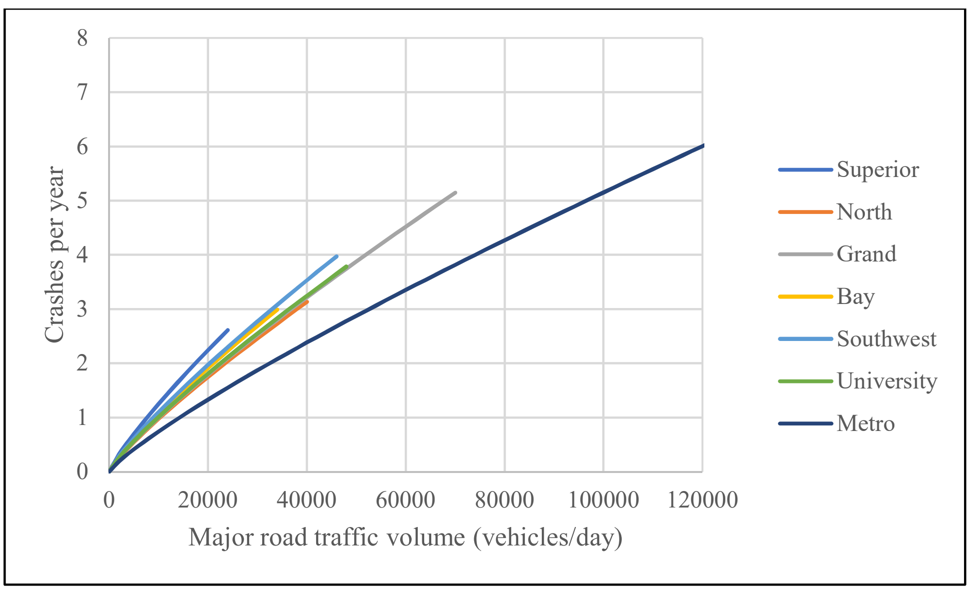


Figure 13. Graphical Form of Fatal and Injury Crash AADT and Regional Indicator SPF for Three-Leg Signalized (4SG) Intersections

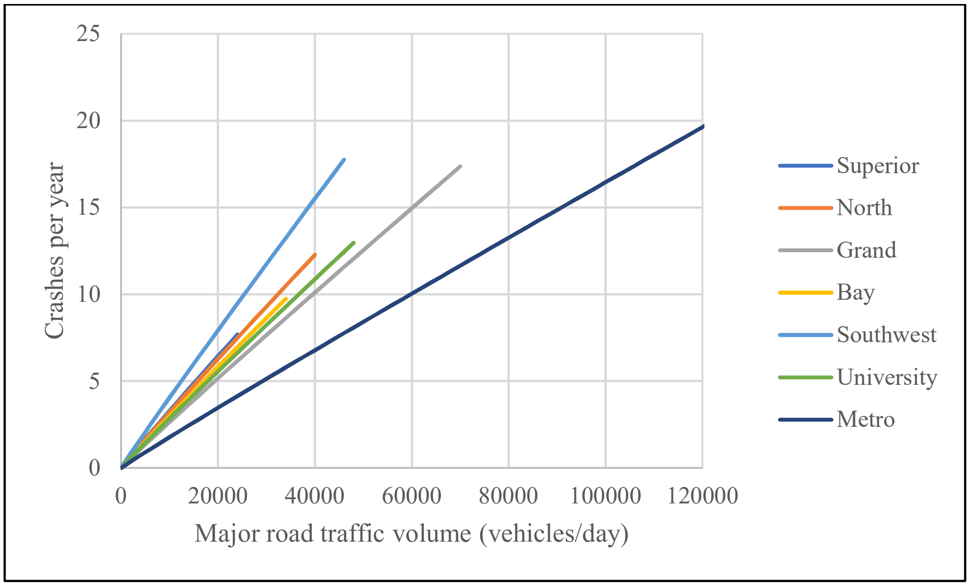


Figure 14. Graphical Form of PDO Crash AADT and Regional Indicator SPF for Three-Leg Signalized (4SG) Intersections

Similar tables and figures were made for each of the other intersection types explored in this study, however, they have been omitted from this document as the models and figures are similar across all facility types. Perhaps the most interesting takeaway from the preceding tables and figures is the additive nature of fatal/injury crashes and PDO crashes relative to total crashes. This suggests that fatal and injury crashes typically occur at locations that experience higher frequencies of PDO crashes. In other words, the more crashes at a site, the more likely one is to be fatal. Moving forward, this suggests that potential countermeasures for reducing the frequency of severe crashes can be identified by analyzing all crashes at an intersection. In some cases, specific crash types may be of particular interest to a researcher or road agency. To this end, Table 9 provides statewide details of the crash type distributions for each of the four site types by severity level (fatal/injury versus property damage only).

Table 9. Statewide Distribution of Crashes by Collision type

Manner of Collision	Proportion of Crashes by Severity Level for Specific Intersection Types							
	3SG		3ST		4SG		4ST	
	FI	PDO	FI	PDO	FI	PDO	FI	PDO
Single Vehicle	0.04	0.05	0.03	0.06	0.02	0.02	0.05	0.06
Rear-end	0.42	0.51	0.28	0.35	0.35	0.45	0.16	0.24
Rear-end Left-turn	0.01	0.02	0.04	0.03	0.01	0.02	0.02	0.02
Rear-end Right-turn	0.01	0.03	0.03	0.04	0.01	0.02	0.01	0.02
Head-on	0.02	0.01	0.03	0.01	0.01	0.01	0.01	0.00
Head-on Left-turn	0.13	0.04	0.11	0.04	0.09	0.05	0.07	0.03
Angle	0.25	0.20	0.30	0.32	0.37	0.25	0.53	0.44
Sideswipe-Same	0.02	0.10	0.03	0.11	0.02	0.12	0.03	0.11
Sideswipe-Opposite	0.01	0.02	0.01	0.01	0.01	0.02	0.01	0.02
Other MV	0.02	0.03	0.01	0.03	0.02	0.03	0.02	0.04
Pedestrian	0.04	0.00	0.05	0.00	0.03	0.00	0.04	0.00
Bicycle	0.03	0.00	0.07	0.01	0.04	0.00	0.05	0.00

2.5.3 Development of Fully Specified SPFs

A variety of approaches could be utilized to develop fully specified SPFs due to the structure of the data set. The data was initially collected and organized such that each observation represented the details for one site for a given year. The simplest methodology to account for this fact is to combine all five years of data for each site into a single observation. For the sake of ease of interpretation, it is beneficial to include the duration of the study as an offset term, which results in scaling of the parameters such that the predictive equation can be applied to estimate the yearly crashes at an intersection. The expression given by Equation 6 and representing the estimated number of crashes for a given observation can thus be expanded as follows:

$$\lambda_i = exp(\beta_0 + \beta_1 X_1 + \cdots + \beta_n X_n + \beta_{duration} X_{duration} + \varepsilon_i). \quad (15)$$

where λ_i represents the five year predicted crash total at a given intersection; ε_i represents the gamma-distributed error term discussed in Chapter 4; $X_1 \ldots X_n$ represent the variables that describe the intersection geometric and operational characteristics; β_0, β_1, … β_n represent estimable parameters; while $X_{duration}$ and $\beta_{duration}$ represent the natural log of the duration (number of years of observations) of the study at a particular intersection and the parameter associated with the number of years of data observed at particular site constrained to equal one, respectively. Making the appropriate substitutions yields the following expression:

$$\lambda_i = exp(\beta_0 + \beta_1 X_1 + \cdots + \beta_n X_n + 1 * ln(duration) + \varepsilon_i). \tag{16}$$

Simple algebra yields the following expression:

$$\lambda_i = duration * exp(\beta_0 + \beta_1 X_1 + \cdots + \beta_n X_n + \varepsilon_i). \tag{17}$$

These steps ensure that the parameter estimates given from the GLM fitting algorithm are in terms of a single year.

Table 10 through Table 13 document the results of the model estimation for the five-year aggregated models. Parameters with a Z-Value of approximately 1 or larger are shown for each model.

2.5.3.1 Fully-Specified SPFs without Regional Indicators Using Five-year Aggregated data

Table 10. Fully-specified SPF for signalized three-leg intersections

Parameter	Estimate	Std. Error	z-Value	p-Value
Intercept	-7.493	1.043	-7.190	<0.001
Natural Log of Major Road AADT	0.634	0.108	5.880	<0.001
Natural Log of Minor Road AADT	0.225	0.035	6.390	<0.001
Major Road Through Lanes	0.053	0.045	1.190	0.235
Right Turns on Minor Road	0.293	0.143	2.040	0.041
RTOR Prohibited	-0.172	0.162	-1.060	0.288
Left Turns on Minor Road	0.188	0.142	1.320	0.186
Major Road Median Presence	-0.473	0.157	-3.000	0.003
Minor Road Median Presence	-0.283	0.170	-1.670	0.095
Major Road One-Way Indicator	-0.684	0.208	-3.290	0.001
Minor Road One-Way Indicator	-0.453	0.180	-2.520	0.012
Major Road Parking Presence	-0.444	0.212	-2.090	0.036
Major Road Speed Limit	0.007	0.007	1.060	0.287
Minor Road Bike Lane Presence	-0.784	0.588	-1.330	0.182
Minor Road Driveways	0.058	0.037	1.540	0.125
Non-motorized Path within 1 Mile	0.258	0.109	2.370	0.018
Lighting Presence	0.135	0.115	1.170	0.241
Bus Stop Presence	-0.169	0.125	-1.350	0.178
Overdispersion parameter	0.367	0.044		

Table 11. Fully-specified SPF for unsignalized three-leg intersections

Parameter	Estimate	Std. Error	Z-Value	p-Value
Intercept	-10.0383	1.542	-6.510	<0.001
Natural Log of Major Road AADT	0.7276	0.135	5.380	<0.001
Natural Log of Minor Road AADT	0.4192	0.083	5.080	<0.001
Right Turns on Major Road	0.7477	0.294	2.540	0.011
Right Turns on Minor Road	0.593	0.315	1.880	0.060
Left Turns on Major Road	0.1731	0.164	1.060	0.291
Minor Road Median Presence	-0.6747	0.389	-1.730	0.083
Major Road Parking Presence	-0.9354	0.386	-2.430	0.015
Major Road Speed Limit	-0.0142	0.012	-1.240	0.215
Major Road Driveways	-0.0617	0.037	-1.660	0.097
Minor Road Driveways	0.1272	0.067	1.910	0.057
Lighting Presence	0.3515	0.201	1.750	0.081
Overdispersion parameter	0.211	0.019		

Table 12. Fully-specified SPF for signalized four-leg intersections

Parameter	Estimate	Std. Error	Z-Value	p-Value
Intercept	-7.305	0.609	-12.000	<0.001
Natural Log of Major Road AADT	0.590	0.065	9.140	<0.001
Natural Log of Minor Road AADT	0.154	0.031	4.950	<0.001
Major Road Through Lanes	0.102	0.029	3.450	0.001
Minor Road Through Lanes	0.233	0.036	6.440	<0.001
Right Turns on Major Road	0.175	0.067	2.620	0.009
Right Turns on Minor Road	0.223	0.064	3.490	<0.001
RTOR Prohibited	-0.309	0.098	-3.160	0.002
Left Turns on Major Road	0.234	0.075	3.140	0.002
Left Turns on Minor Road	0.152	0.072	2.100	0.035
Major Road Median Presence	-0.107	0.114	-0.950	0.344
Minor Road Median Presence	-0.634	0.152	-4.170	<0.001
Major Road One-Way Indicator	0.194	0.098	1.970	0.049
Minor Road One-Way Indicator	0.158	0.104	1.510	0.131
Major Road Parking Presence	-0.300	0.090	-3.320	0.001
Major Road Speed Limit	0.006	0.004	1.660	0.097
Minor Road Bike Lane Presence	0.351	0.192	1.830	0.067
Major Road Driveways	0.015	0.011	1.340	0.180
Non-motorized Path within 1 Mile	0.330	0.059	5.560	<0.001
Lighting Presence	0.358	0.140	2.550	0.011
Overdispersion parameter	1.404	0.179		

Table 13. Fully-specified SPF for unsignalized four-leg intersections

Parameter	Estimate	Std. Error	Z-Value	p-Value
Intercept	-8.699	1.043	-8.340	<0.001
Natural Log of Major Road AADT	0.635	0.100	6.350	<0.001
Natural Log of Minor Road AADT	0.386	0.060	6.430	<0.001
Minor Road Through Lanes	0.367	0.215	1.710	0.087
Right Turns on Minor Road	0.317	0.283	1.120	0.263
Left Turns on Major Road	-0.130	0.111	-1.170	0.243
Left Turns on Minor Road	0.435	0.179	2.430	0.015
Major Road Parking Presence	-0.185	0.153	-1.210	0.225
Major Road Speed Limit	-0.009	0.008	-1.140	0.256
Major Road Bike Lane Presence	-0.509	0.384	-1.330	0.185
Major Road Driveways	0.028	0.030	0.950	0.342
Minor Road Driveways	-0.034	0.027	-1.260	0.208
Non-motorized Path within 1 Mile	0.157	0.112	1.400	0.161
Presence of Flashing Beacons	0.601	0.262	2.300	0.022
Sidewalk Presence	-0.343	0.165	-2.090	0.037
Skew Angle	-0.006	0.004	-1.380	0.167
Overdispersion parameter	0.598	0.063		

Consistent with expectation, the effect of traffic volume has a substantial effect on the expected crashes at an intersection. In all four models, the coefficient for traffic volume on both major and minor roadways was less than one, indicating as traffic volumes increase, the frequency of traffic crashes increases at a decreasing rate. When comparing three-leg intersections to four-leg intersections, the effect of traffic volume attributed to the minor road is typically larger on four-leg intersections, likely due to the higher number of conflicting movements at four-leg intersections.

Higher numbers of through lanes on the major road were shown to be associated with increasing crash frequency on the signalized intersections, although the effect is less pronounced in the three-leg intersections than four-leg intersections (each lane increased the crash frequency by 5 percent and 11 percent, respectively). This effect is potentially attributable to road users crashing while attempting to maneuver their vehicle to turn on to the minor road or onto nearby driveways.

The number of minor road through lanes was associated with increasing crash frequency on the four-leg intersections. Locations with higher numbers of through lanes on the minor road would likely have relatively higher volumes of traffic on the minor road, therefore it makes sense that they would experience higher crash frequencies. The effect of minor road through lanes was most pronounced at unsignalized intersections, where each lane resulted in an increase of 44 percent, compared to 26 percent for the signalized intersections. The effect is likely more pronounced at unsignalized intersections as multiple through lanes could result in vehicles obscuring the view of drivers, making it more challenging to find an appropriate gap to enter the mainline traffic or to cross the major road. Through lanes on the minor road of three-leg intersections represent intersections where only one lane is present for vehicles to enter the major road, so intuitively it makes sense that the effect of these lanes would be insignificant, especially as this scenario would also be heavily associated with low minor road traffic volumes, and thus, fewer conflicting movements.

The presence of dedicated right turn lanes on the major road of three-leg unsignalized and four-leg signalized intersections was shown to be associated with increased crash frequency. This result is initially counter-intuitive, as one would expect that the dedicated turn lanes would remove vehicles from the traffic flow, reducing the likelihood of rear-end collisions as the deceleration of turning vehicles will not impact the though movement traffic. This result is potentially attributable to the fact that right turn lanes service driveways adjacent to the intersections as well as the intersections themselves. For all four facility types, the presence of right turn lanes on the minor road was associated to be with an elevated crash frequency. This effect is likely related to the fact that dedicated turn lanes on minor streets would deliberately be located at intersections with elevated minor road volumes, leading to more conflicts and crashes.

The prohibition of right turns when the traffic signal is red was associated with decreased crash frequency at signalized intersections. This is consistent with what would be expected for this particular crash countermeasure. At three-leg intersections, the observed reduction in crashes was 16 percent, while four-leg intersections with right-turn-on-red prohibition were associated with 27 percent fewer crashes. The discrepancy between the two facility types is probably partially due to the extra leg that vehicles would be turning on and off of.

The presence of dedicated left turn lanes on the major roadway differed by facility type. No effect was significantly present for the three-leg signalized intersections, while crashes were 19 percent higher at four-leg signalized intersections, 26 percent higher for three-leg unsignalized intersections, and 12 percent lower for four-leg unsignalized intersections. It is somewhat surprising that this effect differs so much for each facility type, however, the varying effects may be indicative of the differing circumstances at which dedicated left-turn lanes are provided: to account for high volumes of turning vehicles and to ensure the operational efficiency of the through movements. It is also worth noting that the parameter for left-turn lanes on the major roadway was much more significant for four-leg signalized intersections that for the unsignalized intersections. Intuitively, the elevated crash frequency at the four-leg signalized intersection makes sense, as the left-turn lane specifically present due to traffic patterns at the site including a relatively large volume of left-turning vehicles, resulting in more conflicts. This effect could potentially be more effectively examined if traffic timing information was available to assess if the movement is permissive (left-turning vehicles are allowed to use reasonable gaps in oncoming traffic to make their turn) or protected (left-turning vehicles only perform the turning maneuver when the traffic signal allows). The crash reduction associated with major leg left turn lanes on four-leg unsignalized intersections could be reflective of left-turn lanes helping to minimize the rear-end collisions that would have occurred had the turning vehicles been left in the path of through traffic. Similarly to the four-leg signalized intersections, the increase in crashes at three-leg unsignalized intersections could potentially be due to the traffic volumes of turning vehicles rather than for the operational efficiency of through vehicles.

Unlike the presence of dedicated left-turn lanes on the major road, dedicated left-turns on the minor road were consistently associated with increased traffic crash frequency, with the exception of no effect being determinable for three-leg stop controlled intersections. The effect is particularly large for four-leg unsignalized intersections (55 percent), which is relatively consistent with the other trends regarding lanes on the minor roadway of unsignalized intersections that were observed as a part of this study. Essentially, every lane on the minor roadways is indicative of more vehicles entering the higher-volume roadway resulting in more conflicting movements, and thus, more crashes.

The presence of medians on the major roadway was associated with reductions in crashes of 38 percent on three-leg signalized intersections and 10 percent on four-leg signalized intersections, although it was only marginally significant for the latter. This is partially attributable due to the prevalence indirect left turns through the use of median-turnarounds, or Michigan Lefts, on divided roadways. This effect was tested for through the use of an interaction term between the presence of left-turn lanes and medians, however there was no noteworthy interaction between the two features.

The presence of medians on the minor road was found to be associated with crash reductions of 25 percent for three-leg signalized intersections, 49 percent for three leg unsignalized intersections, and 47 percent for four-leg signalized intersections. This effect may be reflective of low volumes or relatively few conflicting movements at intersections with minor-road medians, or some other effect that is correlated with minor road medians but that is unaccounted for in this particular data set.

Traffic crashes were 50 percent lower at three-leg signalized intersections where the major leg was a one-way street, while crashes were 21 percent higher at similar four-leg signalized intersections. The a priori expectation was that intersections with one-way streets would be associated with fewer crashes, as was observed on the three-leg signalized intersections. The elevated crash frequency for this type of four-leg signalized intersection could be reflective of specific traffic patterns or due to these locations being most prevalent in the densest urban environments. From a human factors standpoint, it is possible that drivers travelling through the intersection are unaccustomed to vehicles turning onto the roadway and travelling in the same direction as they are when the vehicle enters from the left side. Somewhat surprisingly, the effects are very similar for one-way minor roadways.

To varying degrees, the presence of parking on the major roadway was associated with reduced crash frequency at each of the four facility types (36, 61, 26, and 17 percent at three-leg signalized, three-leg unsignalized, four-leg signalized, and four-leg unsignalized, respectively). This is quite possibly the result of a selectivity bias, in that parking would only be allowed where the crash history would suggest that it is relatively safe to do so. This is also potentially another instance where the presence of parking is associated with other safety features that are not accounted for in this data set.

The speed limit of the major roadway was found to be associated with increased crash frequency for signalized intersections (about 3.6 percent for every five miles per hour at three-leg signalized and 3 percent for every five miles per hour at four-leg signalized). For unsignalized intersections, higher major road speed limits were associated with decreased crash frequency (7 percent for every five miles per hour on three-leg intersections and 4.5 percent for four-leg intersections). This result makes sense, as higher speed limits result in increased stopping sight distances at signalized intersections, which may result in dilemma zone related crashes, where drivers are caught in an area where they cannot determine whether they should proceed through the intersection or bring their vehicle to a stop. For unsignalized intersections, the reduced crash frequency is likely again due to selection bias, as speed limits would likely be relatively lower at intersections where substantial volumes of vehicles are entering the major roadway from the minor roadway and higher when the volume of traffic on the minor roadway is relatively low.

The presence of a dedicated bike lane on the major road was shown to be associated with a 40 percent reduction of traffic crashes for four-leg unsignalized intersections. This is likely due to bike lanes being installed where engineering judgement deemed that they could be installed in a relatively safe manner. Additionally, the presence of bike lanes may be indicative of more available right-of-way, and thus, more room for vehicles to maneuver and avoid potential crashes. The presence of dedicated bike lanes on the minor roadway was found to be associated with a 54 percent lower crash frequency at three-leg signalized intersections and a 43 percent higher crash frequency at four-leg signalized intersections. It is hard to say why these have differing signs, although it may be due to relatively low representation within the data, or site-specific characteristics at those sites where the bike lanes are located.

For each driveway within 210 feet of the center of the intersection, traffic crashes decreased by 6 percent at three-leg stop controlled intersections, while increasing by 1.5 and 2.9 percent at four-leg signalized and unsignalized intersections, respectively. It is somewhat intuitive that effect of driveways located near to the intersection would have a different effect on three-leg intersections versus four-leg, as one side of the major roadway would typically be free for development at locations with three-legs, however, it is counterintuitive that more driveways would ever be associated with fewer crashes, unless the relationship between the number of driveways and crash frequency at unsignalized intersections is actually reflecting that locations with high driveway counts have lower traffic volumes and thus fewer crashes the other four-leg unsignalized sites.

Each driveway on the minor roadway was shown to increase traffic crashes at three-leg signalized intersections by 6 percent, crashes at three-leg unsignalized intersections by 14 percent, and decrease crashes at four-leg unsignalized intersections by 3 percent. The results of the three-leg intersections were consistent with expectation, as the larger number of driveways immediately adjacent to the intersection is indicative of a high density of conflicting movements. The result for four-leg unsignalized intersections is potentially to many driveways being for single-family homes, and therefore not reflective of a large volume of turning vehicles.

Perhaps one of the most interesting findings of this study pertains to the association between crash frequency and proximity to recreational non-motorized paths. When located within one mile of a non-motorized path, three-leg signalized (29 percent), four-leg signalized (39 percent), and four-leg unsignalized intersections (17 percent) were all shown to exhibit higher crash frequencies. The nature of this finding is somewhat unclear, with one potential explanation being these intersections generally tend to serve lower volumes of traffic. Similar to the effects shown by the regional indicator variables: parameters that are generally associated with lower volume locations tend to indicate increased crash frequency as crashes tend to increase at a decreasing rate as volume increases. Table 14 shows the average major and minor traffic volumes for each intersection type when a non-motorized path is located within one mile or less versus more than one mile of the intersection, which support the thought process regarding lower volume, particularly for signalized locations.

Table 14. Relationship between traffic volume and non-motorized trail proximity

Facility Type	Non-motorized Trail Proximity	Major AADT	Minor AADT
3SG	1 Mile or Less	19172	3555
3SG	More than 1 Mile	20473	3930
3ST	1 Mile or Less	13232	429
3ST	More than 1 Mile	12958	555
4SG	1 Mile or Less	18616	8318
4SG	More than 1 Mile	22438	9195
4ST	1 Mile or Less	14095	1658
4ST	More than 1 Mile	13411	2003

One of the most counterintuitive results in this study is that the presence of lighting at intersections was shown to be associated with higher crash frequency (14 percent at three-leg signalized, 42 percent at three-leg stop controlled, and 43 percent at four-leg stop controlled intersections). While prior research generally suggests locations with lighting tend to experience fewer crashes, these results are somewhat inconsistent. Within the context of this study, these results may be due to lighting being present primarily at either low-volume locations or sites with a history of excess crashes. In the same vein, the presence of a flashing beacon was found to be associated with an 82 percent increase in crashes on four-leg unsignalized intersections. This is likely reflective of the fact that flashing beacons have been strategically deployed at high crash locations.

The presence of sidewalks were associated with 29 percent fewer crashes at four-leg unsignalized intersections. Given the urban/suburban nature of the data, locations without sidewalks are the minority. This effect is again likely picking up on a volume-related issue at these sites. Four-leg signalized intersections were also shown to have decreasing crash frequency as the skew angle of the intersection increased. Large skew angles are relatively rare in the data set, but it is possible that drivers use more caution at locations with high skew, or that the high skew intersections serve relatively minimal traffic volumes. Bus stop presence near three-leg signalized intersections was shown to be associated with a 16 percent reduction in crashes. This is likely a relic of the data, as bus stops are frequently located on high-volume routes in the most densely populated areas of the state.

2.5.3.2 Fully Specified SPFs with Regional Indicators using five-year aggregated data

Tables 15 through 18 present the SPFs developed on aggregated data and utilizing regional indicator variables. Generally speaking, the inclusion of regional indicator variables had a relatively low impact on the coefficients of other variables, with the primary exception being the traffic volume coefficients.

Table 15. Fully specified three-leg signalized intersection SPF with regional indicators

Parameter	Estimate	Std. Error	Z-Value	p-Value
Intercept	-9.058	1.126	-8.040	<0.001
Natural Log of Major Road AADT	0.737	0.112	6.570	<0.001
Natural Log of Minor Road AADT	0.228	0.034	6.610	<0.001
Major Road Through Lanes	0.076	0.048	1.580	0.114
Right Turns on Minor Road	0.215	0.140	1.530	0.125
RTOR Prohibited	-0.264	0.161	-1.640	0.101
Left Turns on Minor Road	0.214	0.138	1.550	0.122
Major Road Median Presence	-0.513	0.155	-3.320	0.001
Minor Road Median Presence	-0.291	0.166	-1.750	0.079
Major Road One-Way Indicator	-0.380	0.213	-1.780	0.075
Major Road Parking Presence	-0.334	0.213	-1.570	0.116
Major Road Speed Limit	0.010	0.006	1.510	0.131
Minor Road Bike Lane Presence	-1.287	0.617	-2.090	0.037
Non-motorized path within 1 Mile	0.200	0.110	1.820	0.069
Lighting Presence	0.109	0.114	0.960	0.339
Terminal Major Leg	0.450	0.341	1.320	0.187
Superior Region	0.516	0.260	1.990	0.047
North Region	0.672	0.178	3.770	<0.001
Grand Region	0.547	0.177	3.090	0.002
Bay Region	0.441	0.185	2.390	0.017
Southwest Region	0.677	0.165	4.120	<0.001
University Region	0.468	0.167	2.800	0.005
Overdispersion Parameter	0.329	0.041		

Table 16. Fully specified three-leg unsignalized intersection SPF with regional indicators

Parameter	Estimate	Std. Error	Z-Value	p-Value
Intercept	-13.043	1.699	-7.670	<0.001
Natural Log of Major Road AADT	0.921	0.161	5.730	<0.001
Natural Log of Minor Road AADT	0.452	0.082	5.520	<0.001
Minor Road Through Lanes	-0.419	0.315	-1.330	0.183
Right Turns on Major Road	0.897	0.279	3.220	0.001
Major Road Parking Presence	-0.822	0.384	-2.140	0.032
Major Road Driveways	-0.053	0.035	-1.500	0.134
Minor Road Driveways	0.133	0.066	2.030	0.043
Lighting Presence	0.278	0.206	1.350	0.178
Sidewalk Presence	0.266	0.199	1.340	0.180
Bus Stop Presence	0.306	0.252	1.220	0.224
Superior Region	0.960	0.360	2.670	0.008
North Region	0.679	0.344	1.980	0.048
Grand Region	0.533	0.316	1.690	0.092
Bay Region	0.451	0.346	1.310	0.192
Southwest Region	1.124	0.343	3.280	0.001
University Region	0.536	0.331	1.620	0.106
Overdispersion Parameter	1.324	0.172		

Table 17. Fully specified four-leg signalized intersection SPF with regional indicators

Parameter	Estimate	Std. Error	Z-Value	p-Value
Intercept	-8.621	0.648	-13.300	<0.001
Natural Log of Major Road AADT	0.680	0.065	10.480	<0.001
Natural Log of Minor Road AADT	0.175	0.030	5.920	<0.001
Major Road Through Lanes	0.141	0.027	5.290	<0.001
Minor Road Through Lanes	0.216	0.034	6.370	<0.001
Right Turns on Major Road	0.206	0.062	3.340	0.001
Right Turns on Minor Road	0.217	0.061	3.530	<0.001
RTOR Prohibited	-0.352	0.094	-3.740	<0.001
Left Turns on Major Road	0.284	0.069	4.130	<0.001
Left Turns on Minor Road	0.079	0.063	1.260	0.208
Minor Road Median Presence	-0.595	0.144	-4.130	<0.001
Major Road One-Way Indicator	0.196	0.096	2.050	0.041
Major Road Parking Presence	-0.307	0.084	-3.660	<0.001
Minor Road Bike Lane Presence	0.307	0.182	1.690	0.091
Major Road Driveways	0.023	0.012	1.900	0.057
Minor Road Driveways	-0.014	0.013	-1.080	0.281
Non-motorized path within 1 Mile	0.257	0.059	4.340	<0.001
Lighting Presence	0.323	0.130	2.470	0.013
Superior Region	0.526	0.118	4.460	<0.001
North Region	0.554	0.112	4.960	<0.001
Grand Region	0.418	0.102	4.080	<0.001
Bay Region	0.546	0.107	5.120	<0.001
Southwest Region	0.710	0.105	6.730	<0.001
University Region	0.588	0.107	5.500	<0.001
Overdispersion Parameter	0.184	0.017		

Table 18. Fully specified four-leg unsignalized intersection SPF with regional indicators

Parameter	Estimate	Std. Error	Z-Value	p-Value
Intercept	-9.583	1.078	-8.890	<0.001
Natural Log of Major Road AADT	0.609	0.090	6.780	<0.001
Natural Log of Minor Road AADT	0.361	0.061	5.950	<0.001
Minor Road Through Lanes	0.396	0.212	1.870	0.062
Left Turns on Minor Road	0.414	0.173	2.390	0.017
Major Road One-Way Indicator	0.240	0.155	1.550	0.121
Major Road Bike Lane Presence	-0.500	0.391	-1.280	0.201
Major Road Driveways	0.028	0.026	1.090	0.274
Non-motorized path within 1 Mile	0.148	0.114	1.300	0.192
Presence of Flashing Beacons	0.544	0.255	2.130	0.033
Sidewalk Presence	-0.171	0.119	-1.430	0.152
Superior Region	0.461	0.215	2.140	0.032
North Region	0.644	0.197	3.280	0.001
Grand Region	0.706	0.185	3.810	<0.001
Bay Region	0.495	0.189	2.610	0.009
Southwest Region	0.553	0.192	2.890	0.004
University Region	0.583	0.188	3.100	0.002
Overdispersion Parameter	0.566	0.061		

2.5.3.3 Fully-Specified SPFs using Random Intercept Generalized Linear Mixed Models and disaggregated data

For the purpose of comparison, a generalized linear mixed model (GLMM) was estimated for a disaggregate version of each of the four data sets. Rather than aggregating all five years of data at each site into one observation, this approach allows for to the level of the observations to be a site-year. To account for the correlation in crash frequency at intersections over the five years, a site-level random effect is used. The result is a data set with five times the number of observations, however the mean of the crash frequency is therefore five times lower. This comparison provides important insight into the effect of sample mean on the variance of the random intercept and subsequently on the overdispersion parameter for researchers the two methodological approaches. Tables 19 through 22 contain the model results for three-leg signalized intersections, three-leg unsignalized intersections, four-leg signalized intersections, and four-leg unsignalized intersections, respectively, which were estimated using a negative binomial model with a site-level random effect.

Table 19. Fully-specified SPF for signalized three-leg intersections with site-level random effect

Parameter	Estimate	Std. Error	z-Value	p-Value
Intercept	-7.619	1.063	-7.170	<0.001
Natural Log of Major Road AADT	0.647	0.107	6.030	<0.001
Natural Log of Minor Road AADT	0.237	0.036	6.570	<0.001
RTOR Prohibited	-0.195	0.169	-1.150	0.250
Left Turns on Minor Road	0.336	0.115	2.930	0.003
Major Road Median Presence	-0.384	0.157	-2.460	0.014
Minor Road Median Presence	-0.256	0.175	-1.460	0.143
Major Road One-Way Indicator	-0.666	0.210	-3.180	0.002
Minor Road One-Way Indicator	-0.503	0.187	-2.680	0.007
Major Road Parking Presence	-0.461	0.210	-2.200	0.028
Major Road Speed Limit	0.008	0.007	1.110	0.266
Minor Road Bike Lane Presence	-1.317	0.589	-2.240	0.025
Minor Road Driveways	0.046	0.037	1.240	0.215
Non-motorized Path within 1 Mile	0.190	0.095	2.010	0.045
Lighting Presence	0.195	0.119	1.630	0.103
Bus Stop Presence	-0.194	0.130	-1.490	0.135
Overdispersion Parameter	0.039	0.014		
Variance of Random Effect				
σ^2	0.387			

In contrast to the five-year aggregated model, the effect of major road through lanes and major road right turn was not significant. Additionally, the effect for minor road left turn presence and minor road bike lane were much more pronounced. With the exception of those aforementioned parameters, the random intercept model is quite similar to the five-year grouped model. The variance of the random effect was estimated to be 0.387, which that some of the heterogeneity of the data is being explained by the random intercept. The subsequent models suggest that the amount of variation in the data explained by the random intercept is directly correlated to sample mean, which is largely a function of traffic volume.

Table 20. Fully-specified SPF for unsignalized three-leg intersections with site level random effect

Parameter	Estimate	Std. Error	z-Value	p-Value
Intercept	-11.4274	1.6416	-6.96	<0.001
Natural Log of Major Road AADT	0.870	0.147	5.930	<0.001
Natural Log of Minor Road AADT	0.414	0.089	4.630	<0.001
Right Turns on Major Road	0.855	0.318	2.690	0.007
Right Turns on Minor Road	0.631	0.320	1.970	0.048
Minor Road Median Presence	-0.524	0.409	-1.280	0.200
Major Road One-Way Indicator	-0.354	0.244	-1.450	0.147
Major Road Parking Presence	-0.946	0.389	-2.430	0.015
Major Road Speed Limit	-0.024	0.012	-2.070	0.039
Minor Road Bike Lane Presence	1.401	1.200	1.170	0.243
Major Road Driveways	-0.048	0.038	-1.260	0.208
Minor Road Driveways	0.124	0.066	1.880	0.060
Lighting Presence	0.338	0.208	1.630	0.104
Overdispersion Parameter	0.113	0.057		
Variance of Random Effect				
σ^2	1.219			

The random effect model for three-leg unsignalized intersections indicated that there was an association between reduced crash frequency and the major roadway being a one-way as well as a crash increase associated with the presence of a minor road bike lane. The effect of major road left turn lane presence was no longer significantly observable. The model is relatively similar to the five-year grouped model barring those changes. The variance associated with the random effect is larger for this data set than for the others, indicating a potential relationship between mean crash frequency (this data set has the lowest) and site level variation. This suggests that less of the variation of the data is explainable with the captured variables, resulting in more of the data being explained by the random intercept term. The low frequency of crashes is generally attributable to relatively low traffic volume, particularly on the minor road.

Table 21. Fully-specified SPF for signalized four-leg intersections with site-level random effect

Parameter	Estimate	Std. Error	z-Value	p-Value
Intercept	-7.131	0.615	-11.600	<0.001
Natural Log of Major Road AADT	0.537	0.064	8.430	<0.001
Natural Log of Minor Road AADT	0.173	0.033	5.250	<0.001
Major Road Through Lanes	0.121	0.029	4.220	<0.001
Minor Road Through Lanes	0.220	0.036	6.080	<0.001
Right Turns on Major Road	0.228	0.069	3.320	0.001
Right Turns on Minor Road	0.226	0.066	3.410	0.001
RTOR Prohibited	-0.322	0.103	-3.130	0.002
Left Turns on Major Road	0.294	0.074	3.950	<0.001
Left Turns on Minor Road	0.125	0.074	1.700	0.090
Major Road Median Presence	-0.111	0.115	-0.970	0.334
Minor Road Median Presence	-0.746	0.162	-4.620	<0.001
Major Road One-Way Indicator	0.209	0.101	2.060	0.040
Minor Road One-Way Indicator	0.142	0.108	1.320	0.187
Major Road Parking Presence	-0.304	0.094	-3.240	0.001
Major Road Speed Limit	0.006	0.004	1.640	0.101
Minor Road Bike Lane Presence	0.452	0.200	2.260	0.024
Major Road Driveways	0.018	0.012	1.550	0.121
Non-motorized Path within 1 Mile	0.224	0.052	4.350	<0.001
Lighting Presence	0.426	0.149	2.870	0.004
Skew Angle	-0.002	0.002	-1.020	0.307
Bus Stop Presence	-0.071	0.068	-1.030	0.302
Overdispersion Parameter	0.014	0.005		
Variance of Random Effect				
σ^2	0.228			

The random effect model estimated for four-leg signalized intersections Illustrated that the effect of bus stop presence and skew angle were marginally significant, with relatively large p-Values in comparison to most of the other model parameters. Again, the standard deviation associated with the random effects is relatively large, indicating that there is site-to-site variation. This data set, which has the highest mean crash frequency, has the lowest random effect standard deviation, adding to the evidence of a relationship between sample mean and the site-specific variation accounted for by the random effect. This suggests that as sample mean increases, the variables in the data set are better able to explain the heterogeneity present in the data leaving less to be explained by the random intercept.

Table 22. Fully-specified SPF for unsignalized four-leg intersections with site level random effect

Parameter	Estimate	Std. Error	z-Value	p-Value
Intercept	-9.758	1.039	-9.390	<0.001
Natural Log of Major Road AADT	0.655	0.096	6.820	<0.001
Natural Log of Minor Road AADT	0.363	0.061	5.980	<0.001
Minor Road Through Lanes	0.446	0.203	2.200	0.028
Right Turns on Major Road	0.180	0.169	1.060	0.288
Right Turns on Minor Road	0.266	0.253	1.050	0.294
Left Turns on Major Road	-0.171	0.110	-1.550	0.121
Left Turns on Minor Road	0.424	0.187	2.270	0.023
Major Road Bike Lane Presence	-0.641	0.413	-1.550	0.121
Minor Road Driveways	-0.029	0.023	-1.260	0.209
Presence of Flashing Beacons	0.568	0.252	2.250	0.024
Overdispersion Parameter	0.061	0.025		
Variance of Random Effect				
σ^2	0.594			

The random effect model for four-leg unsignalized intersections is most different from the five-year grouped models. While the presence of right turn lanes on the major roadway was shown to be associated with increased crash frequency, no significant effect could be detected for major road driveways, sidewalk presence, major road parking, proximity to a non-motorized path, major road speed limit, and skew angle. The magnitude of the coefficients common to both models was relatively similar. The standard deviation associated with the random effect for this model demonstrates that for this data, the size of the random effect is somewhat proportional to the sample mean of crash frequency.

One important difference between all of the five-year grouped models and the random effect model is the scale of the overdispersion parameter. For the random effects models, the overdispersion parameter is much smaller, indicating that the models are converging on the Poisson model. This is reflective of the fact that a substantial amount of variation in the crash data is due to site-specific characteristics that were not captured in this data set. This observation has important implications in the spatial transferability of SPFs specifically in the application of empirical Bayes methodology, wherein the emphasis placed on the SPF would be relatively larger if the random effect model was used instead of the five-year aggregated model. Because the estimation of the random effects model results in the estimation of site-specific intercept terms, accuracy of crash prediction within the sample is typically very high (and site-specific crash history is considered in the model estimation), therefore application of empirical Bayes methodology is not necessary, however the same cannot be said for crash prediction outside of the sample.

In order to truly gain insight into the predictive capabilities of each of the models discussed in this paper, each model was evaluated in terms of MAD and MSPE by estimating the crash frequency for each site year using the disaggregate data set. Additionally, the mean crash frequency and mean deviance (which does not utilize an absolute value and therefore allows for a determination as to whether a model tends to over- or under-predict) are also shown. The results of the goodness-of-fit calculations are shown in Table 23.

Table 23. Goodness of Fit Metrics for Each Model Type

Metric	Model	Facility Type			
		3SG	4SG	3ST	4ST
	Mean Observed Crashes	3.900	8.410	0.490	1.230
Mean Deviance	AADT Only	0.016	0.103	-0.015	0.014
	AADT with Regional Indicators	0.049	0.061	-0.016	0.008
	Fully-Specified Aggregated	-0.108	-0.489	-0.124	-0.050
	Aggregated with Regional Indicators	0.107	-0.164	-0.081	-0.020
	Aggregated using Site-Specific EB	-0.056	-0.324	-0.072	-0.057
	Random Effects (Coefficients Only)	-0.727	-1.090	-0.257	-0.305
	Random Effects (with Site-specific effect)	-0.181	-0.314	-0.092	-0.065
MAD	AADT Only	2.699	4.461	0.631	1.122
	AADT with Regional Indicators	2.524	4.322	0.616	1.096
	Fully-Specified Aggregated	2.345	3.880	0.570	1.122
	Aggregated with Regional Indicators	2.318	3.610	0.574	1.099
	Aggregated using Site-Specific EB	1.607	2.396	0.464	0.855
	Random Effects (Coefficients Only)	2.340	3.771	0.529	1.071
	Random Effects (with Site-specific effect)	1.480	2.137	0.399	0.766
MSPE	AADT Only	13.321	47.710	1.212	2.966
	AADT with Regional Indicators	11.661	43.493	1.145	2.864
	Fully-Specified Aggregated	10.518	35.081	1.217	3.238
	Aggregated with Regional Indicators	10.009	28.873	1.130	3.012
	Aggregated using Site-Specific EB	4.696	10.770	0.684	1.593
	Random Effects (Coefficients Only)	11.466	33.809	1.317	3.085
	Random Effects (with Site-specific effect)	4.188	8.674	0.579	1.268

As one would logically suspect, the simple, AADT-only models tended to have the highest prediction error. The models which considered regional indicators as well as AADT were slightly more effective at predicting crashes than the AADT-only models. The aggregated, fully-specified models with regional indicators outperform the aggregated, fully-specified model without regional indicators, however it is clear that the geometric variables explain more variation in the data than the regional information. Considering only the parameter estimates obtained from the random effects models, performance is comparable between the random effects and aggregate models, however, when the site-specific intercept terms are used to estimate crashes the random effects framework fits substantially better than the other models, with only the EB approach being comparable. Ultimately, random effects and EB are slightly different means to accomplish the same thing, which is to account for location-specific crash history in the estimation of crashes. Both of these methodologies balance crash history at a specific site with the expected crash frequency based on similar locations, . The similarity between the parameter estimates and goodness-of-fit of the model on aggregate data and the random effect model on disaggregate data without the inclusion of the site-specific intercept term supports the idea that SPFs should be developed using aggregate data rather than disaggregate data with a random effects model in the development of SPFs with the intent of out-of-sample crash prediction. Furthermore, utilizing site-specific EB is more practical for predicting outside of the sample of intersections, as the site-specific intercept terms would not be known.

2.6 Conclusions

A variety of safety performance functions were estimated in this analysis. First, simple, volume-only naïve-pooled negative binomial models were estimated to gain high-level insight as to how Michigan-specific SPFs compare to those presented in the HSM. Second, in order to investigate the usefulness of regional indicators in crash prediction models, a series of naïve-pooled negative binomial models containing only volume and region information were estimated. Next, a series of three types of fully-specified negative binomial SPFs were estimated which took into account a variety of geometric and operational characteristics: five-year aggregated, five-year aggregated with regional indicators, and a five-year disaggregated random intercept.

The purpose of this series of model estimations is multifaceted. First, the models provide documentation that the inclusion of regional indicators is useful in identifying differences in safety performance due to characteristic that are geographically based without inhibiting the potential to make inferences about various geometric characteristics for this data set. Second, the models estimated in this section demonstrate high levels of similarity between the coefficients estimated in the five-year aggregated and five-year disaggregated random intercept framework when the site-specific intercept terms are not included in crash prediction, however the random intercept term greatly outperforms the other models when the site-specific terms are utilized. Finally, the efforts put forth in this study serve to aid researchers in selecting a modeling framework for SPF development, particularly if the end goal is crash prediction as opposed to development of an explanatory model.

CHAPTER 3. SAFETY PERFORMANCE FUNCTIONS FOR PEDESTRIAN AND CYCLIST CRASHES

3.1 Background

Worldwide, approximately 27 percent of all road traffic deaths occur among pedestrians and cyclists (WHO 2013). In the United States, pedestrians accounted for approximately 12 percent of all traffic fatalities (NHTSA 2009A) while pedalcyclist (operators of two-wheel or more human-powered vehicles) fatalities account for slightly less than two percent of all traffic fatalities (NTSAA 2009B). The development of safety performance functions for pedestrians and cyclists poses several challenges to road agencies. The first major issue that must be addressed is the collection of exposure data. Pedestrian and cyclist exposure data is difficult for researchers to obtain, as most road agencies do not collect such information. In order to circumvent this issue, certain studies have utilized methodologies to estimate pedestrian volume. A study utilizing data from Austin, Texas estimated walk-miles traveled to predict crashes at the census tract level (Wang and Kockelman, 2013). A subsequent study utilized the same estimated pedestrian data to develop a combined exposure variable of vehicle volume plus pedestrian volume (Wang, Sharda, and Wang, 2016). Demographic information has been utilized in the development of crash prediction models, specifically population as an exposure measure (Pulugurtha and Sambhara, 2011). Using a variety of influence radii around intersections, the research found that pedestrian crashes generally increase with population, with the most pronounced effect being within 0.25 miles of a signalized intersection.

At the onset of the project that serves as the basis for this research, it was known that pedestrian and cyclist volumes were not readily available within the state of Michigan. Subsequently, a two-faceted approach to investigating pedestrian and cyclist safety was taken, wherein a series of vehicle-volume based models was developed and then improved upon through the incorporation of data from the American Community Survey.

3.2 Literature Review

Various studies have aimed to investigate specific scenarios that affect crashes involving pedestrians and cyclists in lieu of developing count data models, or developing count data models without pedestrian volumes. A recent study conducted in Beijing found that pedestrians frequently underestimate speed and stopping distance at high-speed crossings (Sun et al. 2015). A study of pedestrian crashes in Israel found that investigated crash typology finding that the majority of fatal and injury pedestrian crashes occurred in urban areas, with most crashes in general occurring at non-crosswalk locations (Gitelman et al. 2012). A study of intersection pedestrian crashes in Florida utilized log-linear models to analyze crash frequency without accounting for pedestrian volume as well as an ordered probit model to analyze pedestrian crash injury severity (Lee and Abdel-Aty, 2005). The study proposed the pedestrian exposure measure of total duration of walking trips. Pedestrian crashes in New Orleans were investigated using cluster analysis to identify locations for auditing for pedestrian safety (Tolford, Renne, and Fields; 2014). A cluster analysis in metropolitan Atlanta utilized spatiotemporal clustering and logistic regression to examine pedestrian crashes, finding suburban-style corridors with long blocks and a mix of high-speed state routes and local roads to be particularly dangerous for pedestrians (Dai 2012). Additionally, behavior such as darting into traffic and not crossing at crosswalks were identified as being particularly hazardous. In a similar vein as some of the previously mentioned cluster analysis, some approaches involve the use of heat maps or other forms of geo-spatial analysis (Kim 2009).

Cyclist crashes in Brussels were evaluated using logistic and autologistic regression models as well as a series of Bayesian logistic regression models (Vandenbulcke, Thomas, and Int Panis, 2014). Tram tracks, proximity to public administration buildings and bridges without cycling facilities were found to be more dangerous for cyclists, while contraflow cycling (bicycle lanes in both directions on a one-way street) was shown to be associated with decreased crash likelihood.

A project focused on developing network screening criteria for pedestrian and cyclist crashes in Portland, Oregon utilized a subjective risk factor identification process in conjunction with spatial analysis in ArcGIS to identify corridors with potential for safety improvement (Bergh and Ray 2014, Bergh et al 2015). Kohonen neural networks were utilized in an Israeli study to identify five types of crashes that were particularly dangerous: urban elderly, pedestrian with two-wheel, rural night, youngsters at night, and rural children (Prato, Gitelman, and Bekhor 2012).

In some cases, counts of pedestrian and cyclist data are available or can be reasonably well estimated. A report from California outlined the process of using log-linear least squares regression models to estimate pedestrian volume based on land use, characteristics of the adjacent roadway, and socioeconomic characteristics (Grembek et al 2014). Research from Montreal modelled cyclist crashes at 647 signalized intersections using a two-equation Bayesian approach to simultaneously model injury frequency and activity (Strauss, Miranda-Moreno, and Morency; 2013). Vehicular turning movement volumes, bus stop presence and crosswalk length were found to increase injury frequency, while the presence of a median was associated with decreased frequency. In terms of cyclist activity, nearby employment, schools, metro stations, mixed land use, amount of bike facilities, and commercial land use were all associated with increased activity, while fewer cyclists were observed at three-approach intersections. The bivariate Bayesian approach was utilized in subsequent research which was expounded upon to estimate models for pedestrian, cyclist, and vehicular crashes. (Strauss, Miranda-Moreno, and Morency; 2014). Turning movements were found to be a significant danger to cyclists, while driveway density was found to be dangerous to pedestrians. Further research utilizing Montreal as a testbed has investigated methodology to incorporate modern technology such as smartphone GPS data with actual bicycle counts to develop predictive models for bicycle volume (Strauss, Miranda-Moreno, and Morency, 2015). A study based in Hamilton, Ontario used a shortest-distance algorithm to predict child-pedestrian activity (Bennet and Yiannakoulias, 2015). This volume estimation coupled with a case-control study design was utilized to develop conditional logistic regression models to examine what roadway characteristics were associated with intersection and mid-block crashes. The study suggests that route choice combined with behavior play a critical role in child pedestrian crashes.

3.3 Data Description and Data Collection

The data set utilized in this study represents a subset of the data collected in Chapter 2. The data in this particular analysis represent only those locations that have sidewalks. This subset was chosen as essentially zero crashes (pedestrian or cyclist) occurred at locations that did not have sidewalks. This suggests that there is a link between the roadway characteristics associated with pedestrian and cyclist crashes. In addition to generating a subset from the data, census-tract level information from the American community survey was associated to every intersection. Intersections which fall on the border of multiple census tracts were given averaged information.

3.3.1 American Community Survey Data

In order to gain further insight into the factors affecting the frequency of pedestrian and cyclist crashes, commuter information was obtained from the American Community Survey (ACS). The ACS is a continually administered survey conducted by the United States Census Bureau and collects a variety of information that had previously been collected in the decennial census.

Approximately 1 in 38 households per year is invited to be a part of the American Community Survey (ACS). Respondents complete a questionnaire via paper or internet and submit it to the U.S. Census Bureau. While the data is available at several levels of granularity, the level utilized in this study was that of the census tract. Census tracts are somewhat proportional to the density of people, with an optimally sized tract containing 4,000 people (census.gov). A map of Michigan census tracts and intersections utilized in this study is shown in Figure 15.

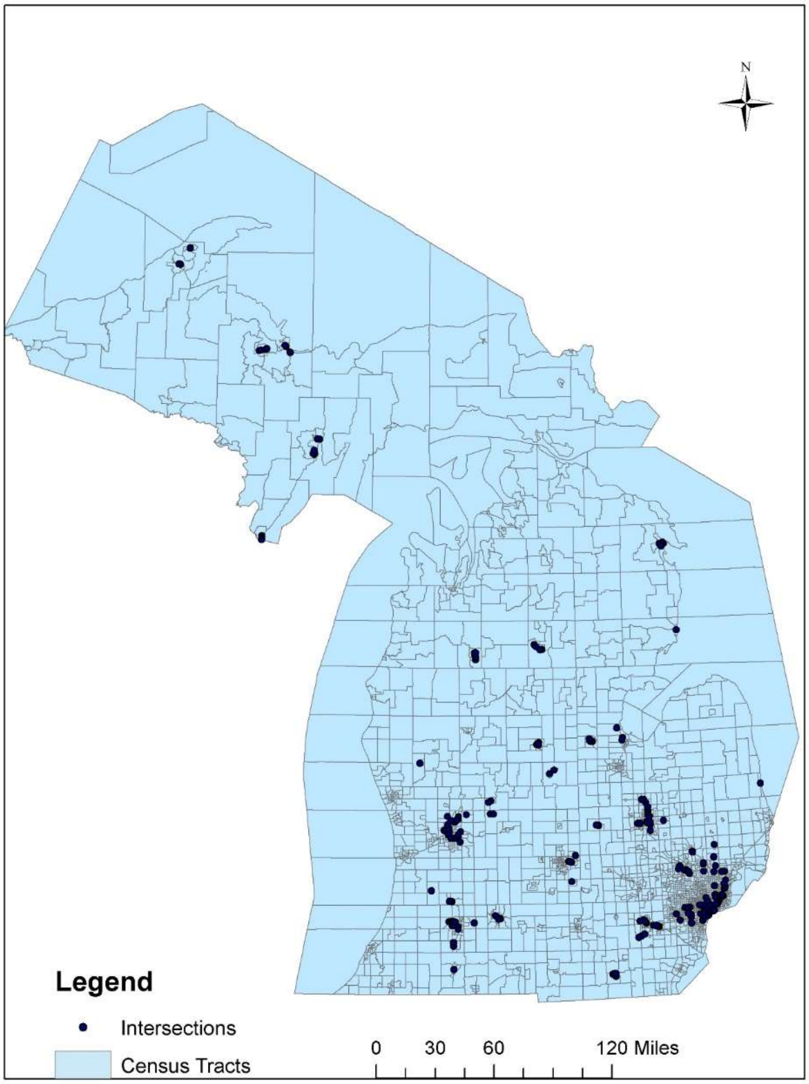


Figure 15. Michigan Census Tracts and Study Intersections

While Figure 15 demonstrates the general size of census tracts across Michigan, it is difficult to make out the tracts in the urban areas, where the tracts have a finer level of granularity. To this end, Figure 16 presents a map which is zoomed to the metro-Detroit area.

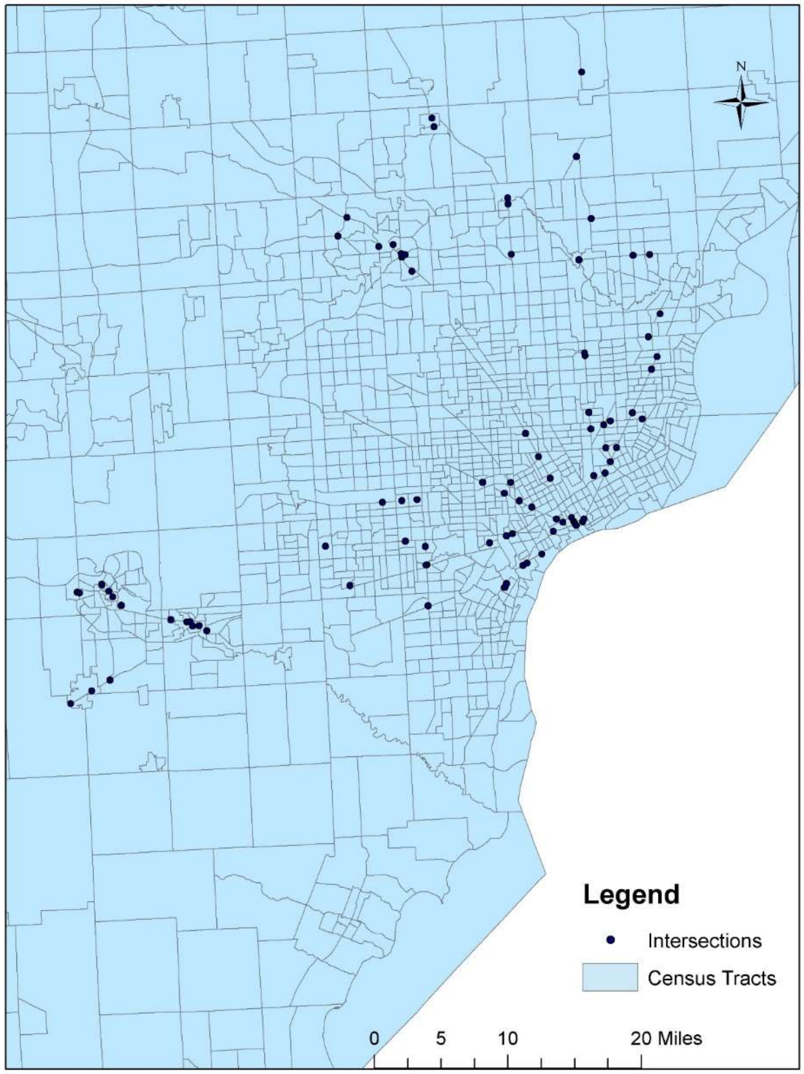


Figure 16. Metro Detroit Census Tracts and Study Intersections

Descriptive statistics for the three-leg and four-leg data used in this analysis are given in Table 24 and Table 25, respectively.

Table 24. Three-leg signalized intersection pedestrian and bicycle crash data set (n=156)

Characteristic	Average	Std. .Dev.	Minimum	Maximum
Major Road Traffic Volume	21196.33	10254.53	4625.00	61372.00
Minor Road Traffic Volume	3920.92	5466.97	48.50	42330.00
Major Road Through Lanes	3.69	1.35	0.00	10.00
Major Road Right Turn Lanes	0.42	0.55	0.00	2.00
Major Road Left Turn Lanes	0.99	0.78	0.00	2.00
Minor Road Through Lanes	0.48	0.79	0.00	4.00
Minor Road Right Turn Lanes	0.70	0.56	0.00	2.00
Minor Road Left Turn Lanes	0.70	0.62	0.00	3.00
Skew Angle	9.91	14.14	0.00	64.08
Lighting Presence	0.79	0.41	0.00	1.00
Right Turn On Red Prohibition	0.12	0.32	0.00	1.00
Major Road Driveway Count	2.53	2.16	0.00	9.00
Minor Road Driveway Count	1.50	1.48	0.00	7.00
Major Road Bike Lanes	0.01	0.11	0.00	1.00
Minor Road Bike Lanes	0.01	0.08	0.00	1.00
Major Road Bus Stop Presence	0.35	0.48	0.00	1.00
Minor Rad Bus Stop Presence	0.07	0.26	0.00	1.00
Major Road Parking	0.12	0.32	0.00	1.00
Minor Road Parking	0.21	0.40	0.00	1.00
Major Road Median Presence	0.10	0.29	0.00	1.00
Major Road Median Width	1.93	7.55	0.00	69.85
Minor Road Median Presence	0.10	0.30	0.00	1.00
Minor Road Median Width	1.66	5.92	0.00	37.67
Superior Region	0.04	0.19	0.00	1.00
North Region	0.10	0.30	0.00	1.00
Grand Region	0.14	0.35	0.00	1.00
Bay Region	0.10	0.29	0.00	1.00
Southwest Region	0.17	0.37	0.00	1.00
University Region	0.15	0.35	0.00	1.00
Metro Region	0.31	0.46	0.00	1.00
Major Road Speed Limit	38.59	7.59	25.00	55.00
One-way Major Road	0.11	0.31	0.00	1.00
One-way Minor Road	0.08	0.28	0.00	1.00
Terminal Major Leg	0.03	0.18	0.00	1.00
Population Density	2358.67	1917.14	102.33	12286.88
Bicycle Commuter Density	14.00	42.50	0.00	349.59
Pedestrian Commuter Density	72.10	229.63	0.00	2331.88
Non-motorized Commuter Density	86.10	265.26	0.00	2681.47
Average Median Household Income	41169.47	19546.23	6765.50	143659.00
Parking	0.24	0.43	0.00	1.00
Hiking trail within 1 Mile	0.03	0.18	0.00	1.00
Biking Trail Within 1 Mile	0.00	0.00	0.00	0.00
Median Age	36.54	7.07	15.55	51.30
Total Pedestrian Crashes	0.24	0.58	0.00	3.00
Total Bicycle Crashes	0.27	0.60	0.00	4.00
Total Non-motorized Crashes	0.51	0.83	0.00	4.00

Table 25. Four-leg signalized intersection pedestrian and bicycle descriptive statistics (n=300)

Characteristic	Average	Std. Dev.	Minimum	Maximum
Major Road Traffic Volume	21705.09	15383.22	4265.00	118771.00
Minor Road Traffic Volume	9070.39	7535.27	94.00	47926.00
Major Road Through Lanes	3.63	1.23	1.00	10.00
Major Road Right Turn Lanes	0.48	0.74	0.00	2.00
Major Road Left Turn Lanes	1.29	0.94	0.00	4.00
Minor Road Through Lanes	2.68	1.10	0.00	8.00
Minor Road Right Turn Lanes	0.44	0.70	0.00	2.00
Minor Road Left Turn Lanes	1.14	0.99	0.00	4.00
Skew Angle	9.66	14.53	0.00	61.04
Lighting Presence	0.98	0.15	0.00	1.00
Right Turn On Red Prohibition	0.90	0.30	0.00	1.00
Major Road Driveway Count	3.73	2.70	0.00	13.00
Minor Road Driveway Count	4.00	2.64	0.00	14.00
Major Road Bike Lanes	0.03	0.18	0.00	1.00
Minor Road Bike Lanes	0.02	0.15	0.00	1.00
Major Road Bus Stop Presence	0.35	0.48	0.00	1.00
Minor Road Bus Stop Presence	0.22	0.42	0.00	1.00
Major Road Parking	0.16	0.37	0.00	1.00
Minor Road Parking	0.18	0.38	0.00	1.00
Major Road Median Presence	0.10	0.30	0.00	1.00
Major Road Median Width	3.64	13.26	0.00	93.80
Minor Road Median Presence	0.04	0.19	0.00	1.00
Minor Road Median Width	1.42	8.69	0.00	72.25
Superior Region	0.12	0.32	0.00	1.00
North Region	0.13	0.33	0.00	1.00
Grand Region	0.14	0.34	0.00	1.00
Bay Region	0.15	0.35	0.00	1.00
Southwest Region	0.15	0.36	0.00	1.00
University Region	0.16	0.36	0.00	1.00
Metro Region	0.16	0.37	0.00	1.00
Major Road Speed Limit	37.17	8.53	25.00	70.00
One-way Major Road	0.13	0.34	0.00	1.00
One-way Minor Road	0.12	0.33	0.00	1.00
Population Density	2476.28	2004.74	34.25	16474.43
Bicycle Commuter Density	14.96	37.31	0.00	306.15
Pedestrian Commuter Density	77.55	203.76	0.00	2514.60
Nonmotorized Commuter Density	92.50	233.25	0.00	2820.75
Average Median Household Income	37332.33	15105.37	9795.00	100080.00
Parking	0.24	0.43	0.00	1.00
Hiking trail within 1 Mile	0.04	0.20	0.00	1.00
Biking Trail Within 1 Mile	0.03	0.18	0.00	1.00
Median Age	36.73	6.86	19.30	56.30
Total Pedestrian Crashes	0.41	0.76	0.00	4.00
Total Bicycle Crashes	0.53	0.96	0.00	8.00
Total Non-motorized Crashes	0.95	1.38	0.00	10.00

Three variables were identified as exposure measures for this analysis: pedestrian commuter density, bicyclists commuter density, and the sum of those densities. In some cases, nobody within a census tract claimed to walk or cycle to work. Since a natural log transformation was going to be used on the data, the data was uniformly translated by adding a value of 1 person per square mile to each density. Figures 17 through 19 present pedestrian, cyclist, and non-motorized crashes versus there appropriate commuter type for three-leg signalized intersections. Three-leg intersections have a lower average number of crashes than their four-leg counterparts (shown in Figures 20 through 22). This is due to the fact that there is one fewer exposure area in which a collision can occur. Each figure is shown with a linear equation and a R^2 value to indicate the general relationship between exposure and crash frequency.

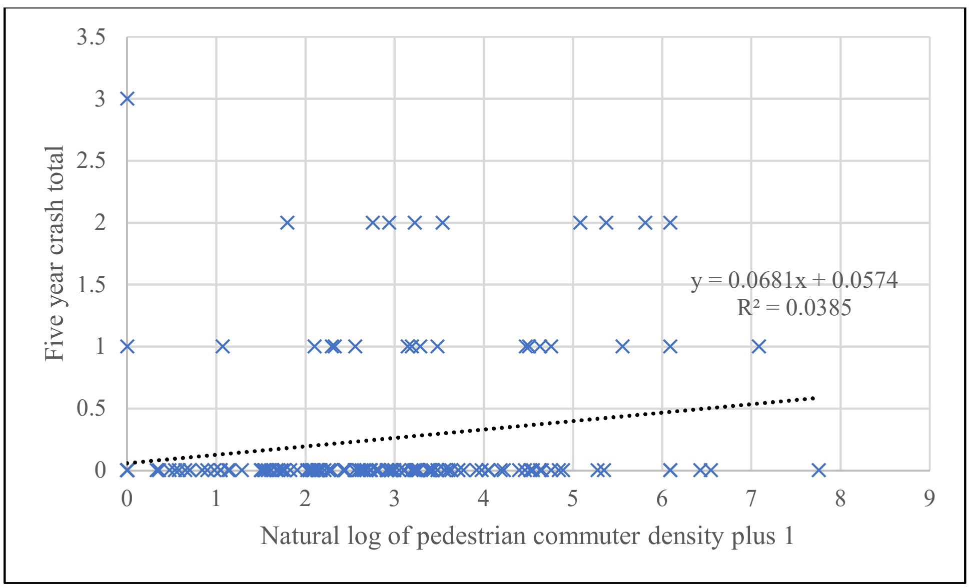


Figure 17. Pedestrian crashes versus pedestrian commuter density at three-leg signalized intersections

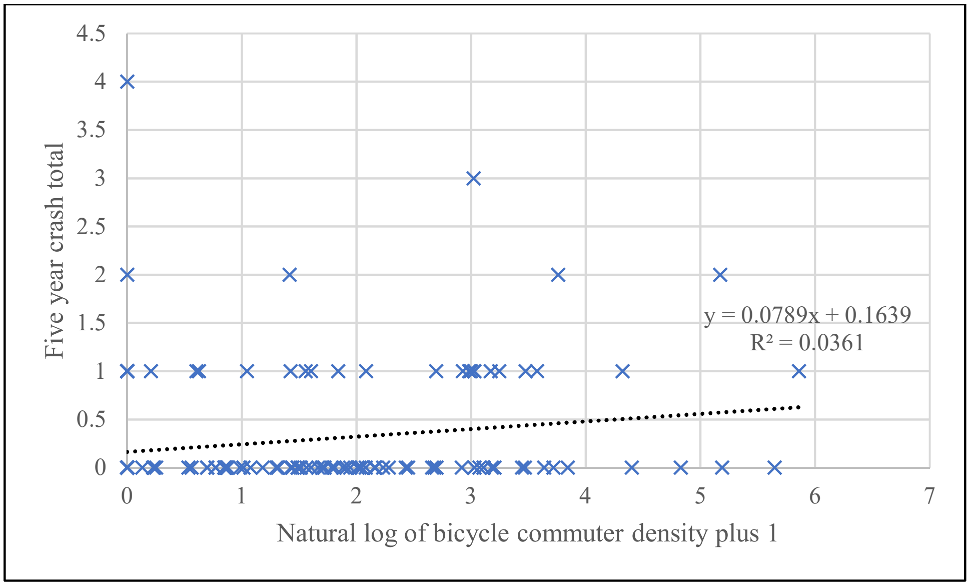


Figure 18. Cyclist crash total versus cyclist commuter density at three-leg signalized intersections

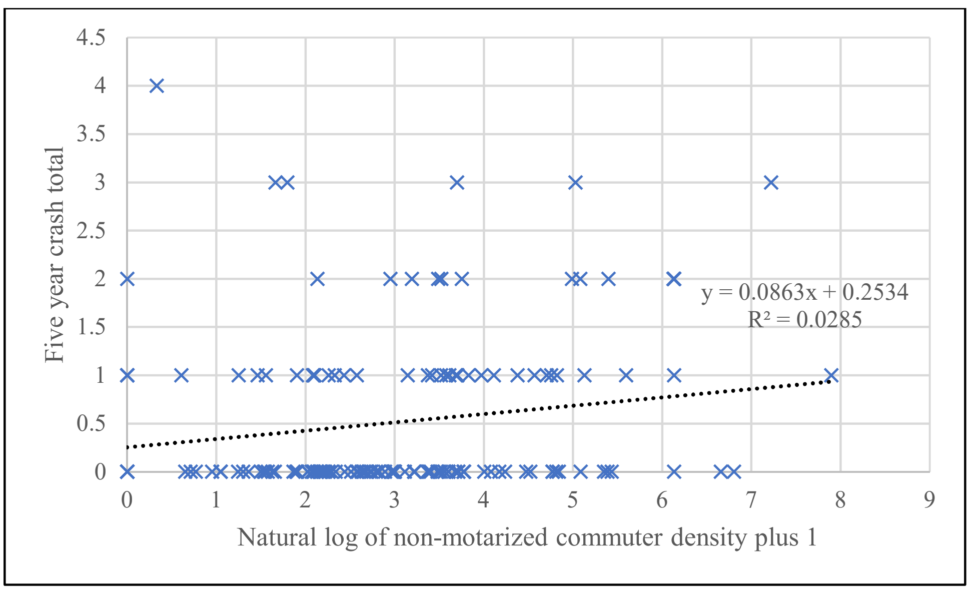


Figure 19. Non-motorized crash total versus non-motorized commuter density at three-leg signalized intersections

The low R^2 values associated with the trend-lines are generally reflective of the fact that factors other than non-motorized commuter volume impact the frequency of non-motorized crashes at intersections. This is further illustrated in figures 20 through 22.

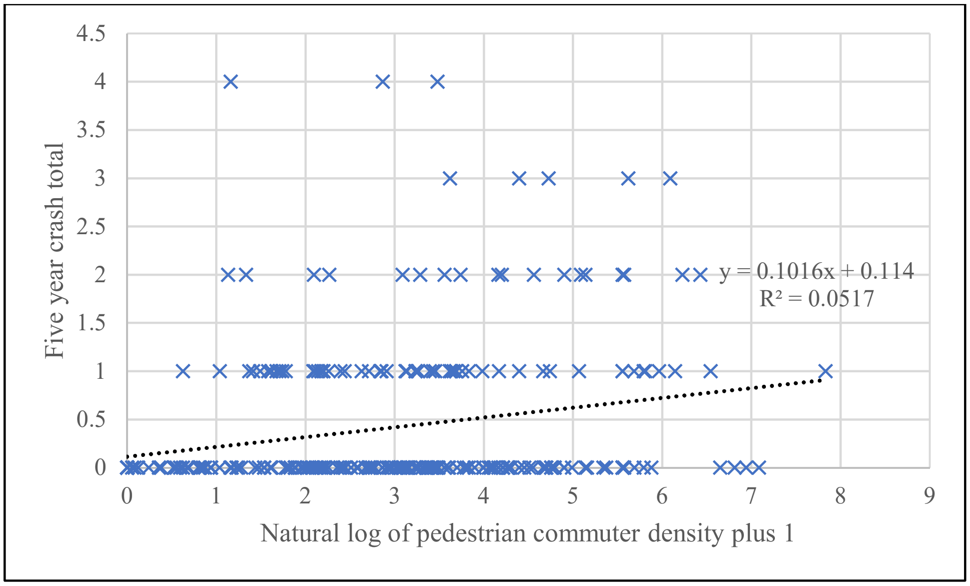


Figure 20. Pedestrian crashes versus pedestrian commuter density at four-leg signalized intersections

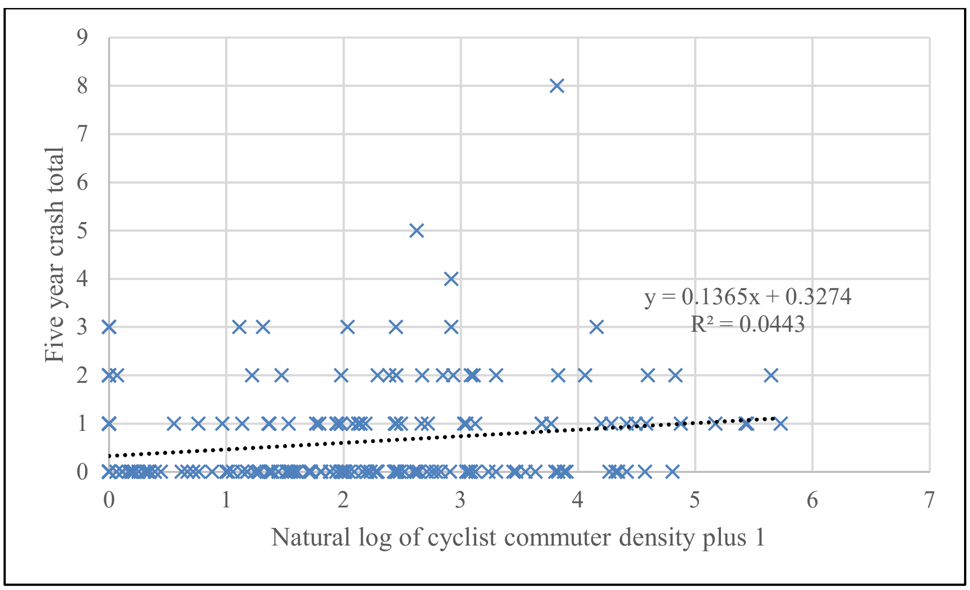


Figure 21. Cyclist crashes versus cyclist commuter density at four-leg signalized intersections

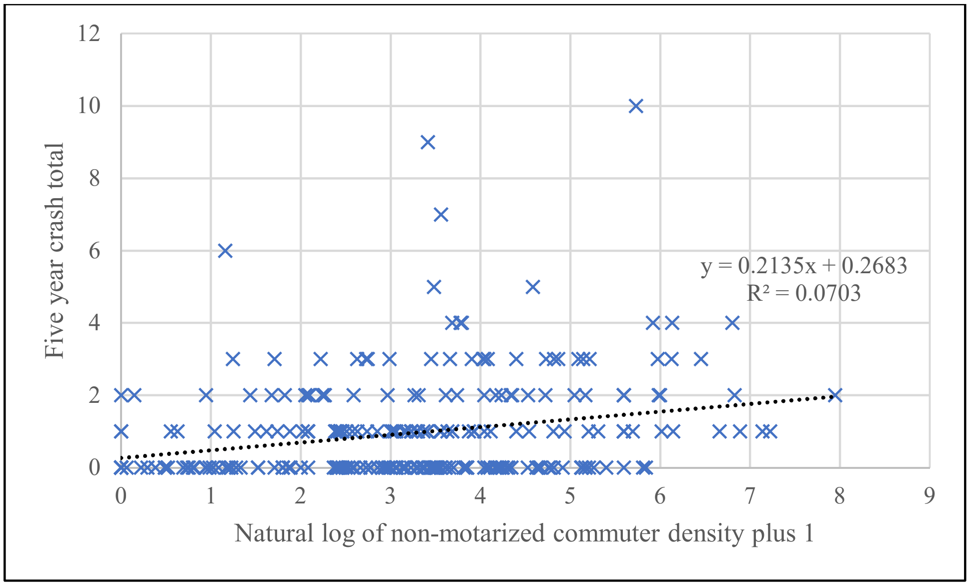


Figure 22. Non-motorized crashes versus non-motorized commuter density at four-leg signalized intersections

3.4 Methodology

These SPFs take the form of generalized linear models. As crash data are comprised of non-negative integers, traditional regression techniques (e.g., ordinary least-squares) are generally not appropriate. Given the nature of such data, the Poisson distribution has been shown to provide a better fit and has been used widely to model crash frequency data. In the Poisson model, the probability of intersection i experiencing y_i crashes during a one-year period is given by Equation 16,

$$P(y_i) = \frac{EXP(-\lambda_i)\lambda_i^{y_i}}{y_i!} \tag{16}$$

where $P(y_i)$ is probability of intersection i experiencing y_i crashes and λ_i is the Poisson parameter for intersection i, which is equal to the segments expected number of crashes per year, $E[y_i]$. Poisson models are estimated by specifying the Poisson parameter λ_i (the expected number of crashes per period) as a function of explanatory variables, the most common functional form being given by Equation 17,

$$\lambda_i = exp(\beta X_i) \tag{17}$$

where X_i is a vector of explanatory variables and β is a vector of estimable parameters.

A limitation of this model is the underlying assumption of the Poisson distribution that the variance is equal to the mean. As such, the model cannot handle overdispersion wherein the variance is greater than the mean. Overdispersion is common in crash data and may be caused by data clustering, unaccounted temporal correlation, model misspecification, or ultimately by the nature of the crash data, which are the product of Bernoulli trials with unequal probability of events (Lord 2006). Overdispersion is generally accommodated through the use of negative binomial models (also referred to as Poisson-gamma models).

The negative binomial model is derived by rewriting the Poisson parameter for each intersection as shown in Equation 18,

$$\lambda_i = exp(\beta X_i + \varepsilon_i) \tag{18}$$

where $EXP\ (\varepsilon_i)$ is a gamma-distributed error term with mean 1 and variance α. The addition of this term allows the variance to differ from the mean as shown in Equation 19:

$$VAR[y_i] = E[y_i] + \alpha E[y_i]^2 \tag{19}$$

The negative binomial model is preferred over the Poisson model since the latter cannot handle overdispersion and, as such, may lead to biased parameter estimates (Lord and Park 2008). Consequently, the *HSM* recommends using the negative binomial model for the development of SPFs.

If the overdispersion parameter (α) is equal to zero, the negative binomial reduces to the Poisson model. Estimation of λ_i can be conducted through standard maximum likelihood procedures. While alternatives, such as the Conway-Maxwell model, have the advantage of accommodating both overdispersion and underdispersion (where the variance is less than the mean) (Lord and Mannering 2010), the negative binomial model remains the standard in SPF development.

The predictive ability of each of the models developed as a part of this research was o assessed using Mean Absolute Deviance (MAD) and Mean Squared Predictive Error (MSPE) (Oh et al. 2003), which are shown in Equations 20 and 21, respectively:

$$MAD = \frac{1}{n} \Sigma_{i=1}^{n} |\hat{u}_i - y_i| \tag{20}$$

and

$$MSPE = \frac{1}{n} \Sigma_{i=1}^{n} (\hat{u}_i - y_i)^2 \tag{21}$$

where:

$n =$ *the number of observations,*
$i =$ *the ith observation,*
$\hat{y}_i =$ *the predicted value of the ith observation, and*
$y =$ *the observed value of the ith observation.*

3.5 Results and Discussion

3.5.1 Preliminary Analysis of Pedestrian and Cyclist Crash Frequency

To gain a fundamental understanding of frequency of pedestrian and cyclist traffic crashes in the state of Michigan, simple SPFs were developed which only accounted for major and minor road traffic volumes. These models were developed to predict total, fatal-injury, and property damage only (PDO) crashes at each of the four primary intersection types (three-leg stop-controlled, three-leg signalized, four-leg stop-controlled, and four-leg signalized) as shown in Table 26 and Table 27.

Table 26: Michigan Specific AADT Only Pedestrian Crash Models

Severity	Intersection Types	Intercept (a)	AADTmaj (b)	AADTmin (c)	Overdispersion factor (k)
Total	3ST	-15.512	0.765	0.385	2.143
	3SG	-9.044	0.402*	0.187	1.057
	4ST	-11.613	0.547	0.269	2.254
	4SG	-7.578	0.364	0.173	0.959
FI	3ST	-15.099	0.742	0.338	1.000
	3SG	-9.223	0.418*	0.182*	1.354
	4ST	-11.52	0.529	0.271	2.712
	4SG	-7.583	0.366	0.157	0.779
PDO	3ST	-20.711	0.886	0.661	1.168E-13
	3SG	-10.221	0.158*	0.283*	1.431E-16
	4ST	-16.547	0.793*	0.247*	<0.001
	4SG	-10.535	0.316	0.311	0.977

*The variable was not significant at 95% confidence interval

Table 27: Michigan Specific AADT Only Bicycle Crash Models

	Intersection Types	Intercept (a)	AADTmaj (b)	AADTmin (c)	Overdispersion factor (k)
Total	3ST	-14.744	0.778	0.394	1.214
	3SG	-11.092	0.575	0.232	1.000
	4ST	-11.173	0.618	0.188	1.184
	4SG	-6.958	0.256	0.227	0.884
FI	3ST	-15.567	0.873	0.353	0.939
	3SG	-10.889	0.551	0.204	1.000
	4ST	-11.555	0.659	0.157	0.083
	4SG	-7.834	0.340	0.203	0.702
PDO	3ST	-13.646	0.340*	0.591	1.648E-07
	3SG	-14.18	0.654*	0.331*	7.56E-11
	4ST	-11.718	0.408*	0.313	1.000
	4SG	-6.087	-0.072*	0.323	0.749

*The variable was not significant at 95% confidence interval

In contrast to the models discussed in Chapter 2, the pedestrian- and bicycle-specific SPFs included AADT and crash data for the entire population of intersection locations. This was due to the fact that the study intersections included a relatively small number of such crashes as well as the fact exposure measures for pedestrians and cyclists were not initially available at the onset of the study, which inhibited the ability to estimate detailed models for non-motorized users.

Each of the models show that crashes increase with respect to major road and minor road traffic volumes. However, even in the highest volume cases, intersections are generally expected to experience only a fraction of a crash per year. In any case, these models provide a general starting point for pedestrian and bicycle safety analyses.

Another point worth noting is that most of the parameters in the PDO models are not statistically significant. This is reflective of the limited number of police-reported pedestrian and bicycle crashes that involve no injuries. Crashes are generally reported if an injury is sustained or, alternately, if more than $1500 in property damage is sustained as a result of the crash. In pedestrian- and bicycle-involved crashes, it is generally unlikely that property damage beyond this level would be sustained without an injury resulting to the non-motorized user. While there are likely to be a significant number of crashes in which no injuries, nor property damage beyond this threshold are sustained, such crashes are not reported by practice.

3.5.2 Detailed Analysis of Pedestrian and Cyclist Crash Frequency

While the models discussed in the previous section provide general insight as to the relationship between traffic crashes involving pedestrians and cyclists and vehicular volume, they do not provide insight into the relationship between crash frequency and geometric characteristics. To gain further insight into the factors affecting pedestrian and cyclist crash frequency, data from the ACS was utilized as a substitute for observational volume data. Substantial effort was made to develop separate models for signalized and unsignalized intersections, however, due to the low frequency of pedestrian and cyclist crashes on unsignalized facilities, this was not possible. Of the original sites collected for SPF development, a subset did not have sidewalks present. Initially, this seemed like a variable of interest for the development of crash prediction models for pedestrian and cyclist crashes, however, further investigation revealed that sites with no sidewalks had virtually no pedestrian or cyclist crashes, and therefore, were removed from the analysis. The following subsections present a total of eight types of models for crashes at signalized intersections. The results of this analysis will provide an insight to the relationship between pedestrian and bicycle crashes as well as a framework for which other researchers can investigate pedestrian and bicyclist crashes.

3.5.2.1 Models for pedestrian crashes by signalized intersection type

Table 28 illustrates the model results for pedestrian crashes at three-leg signalized intersections, Table 29 contains the model results for pedestrian crashes at four-leg intersections, while Table 30 contains a joint intersection model for pedestrian crashes where both three-leg and four-leg intersections were analyzed using the same model, with the two intersection types being differentiated using an indicator variable.

Table 28. Pedestrian crashes at three-leg signalized intersections

Parameter	Estimate	Std. Error	z-Value	p-Value
Intercept	-14.470	5.670	-2.550	0.011
Natural log of walking commuter density	0.168	0.119	1.410	0.158
Natural log of major road traffic volume	1.220	0.546	2.240	0.025
Natural log of minor road traffic volume	0.273	0.155	1.760	0.078
Census tract median income ($10,000s)	-0.286	0.165	-1.730	0.083
Major leg one-way	-1.180	0.740	-1.600	0.111
Major road median presence	-1.240	0.904	-1.370	0.170
Skew angle	-0.029	0.019	-1.520	0.129
Census tract median age	-0.049	0.035	-1.410	0.160
Bay Region	-0.969	0.827	-1.170	0.241
Overdispersion Parameter	0.957	0.700		

Table 29. Pedestrian crashes at four-leg signalized intersections

Parameter	Estimate	Std. Error	z-Value	p-Value
Intercept	-8.290	2.051	-4.041	<0.001
Natural log of walking commuter density	0.159	0.078	2.030	0.042
Natural log of major road traffic volume	0.647	0.207	3.122	0.002
Natural log of minor road traffic volume	0.034	0.102	0.334	0.738
Census tract median income ($10,000s)	-0.129	0.090	-1.430	0.153
Major road median presence	-0.581	0.359	-1.617	0.106
Skew angle	-0.016	0.008	-1.935	0.053
Census tract median age	-0.027	0.019	-1.442	0.149
Minor road parking	0.539	0.245	2.198	0.028
Grand Region	0.759	0.272	2.786	0.005
Bay Region	-0.593	0.317	-1.874	0.061
Overdispersion parameter	0.327	0.229		

Table 30. Joint model for pedestrian crashes at signalized intersections

Parameter	Estimate	Std. Error	z-Value	p-Value
Intercept	-8.226	1.936	-4.249	<0.001
Natural log of walking commuter density	0.119	0.064	1.877	0.060
Natural log of major road traffic volume	0.655	0.190	3.450	0.001
Natural log of minor road traffic volume	0.105	0.082	1.289	0.197
Three-leg intersection	-0.440	0.235	-1.874	0.061
Census tract median income ($10,000s)	-0.174	0.078	-2.221	0.026
Major leg one-way	-0.411	0.275	-1.494	0.135
Minor road median presence	-0.635	0.525	-1.210	0.226
Major road median presence	-0.530	0.330	-1.605	0.108
Skew angle	-0.019	0.008	-2.488	0.013
Census tract median age	-0.037	0.016	-2.295	0.022
Minor road parking	0.440	0.220	2.000	0.045
Grand Region	0.551	0.238	2.319	0.020
Bay Region	-0.410	0.310	-1.320	0.187
Overdispersion parameter	0.495	0.239		

Pedestrian crashes at signalized intersections tend to increase as the density of walking commuters in the census tract of the intersection increase, as well as traffic volume at the intersection increases. These three characteristics are exposure measures indicating the amount of road users at a given location, so their relationship with crash frequency makes sense. Three-leg intersections experience fewer crashes than four-leg intersections primarily due to having fewer conflict points between pedestrians and vehicles.

As the median income of the census tract of the intersection increased, crash frequency tended to decrease. This is likely capturing the fact that more people travel by automobile than by foot in higher income areas, or potentially indicative of infrastructure that is less auto-centric. As the median age of the census tract of the intersection increased, traffic crashes tended to decrease, which is likely indicative of fewer pedestrians.

The three-leg and joint intersection models suggest that one-way major roads are associated with fewer pedestrian crashes which is again due to fewer conflicts between pedestrians and vehicles, specifically turning vehicles. Medians on the major street were shown to be associated with fewer pedestrian crashes in all three models, while minor street medians were shown to be associated with fewer crashes in only the joint model. Intuitively, one would expect that medians would reduce crashes by allowing pedestrians to focus on one direction of traffic at a time while crossing the street and allow slower moving pedestrians or people who enter the crosswalk late in the phase and who may not be able to cross the intersection in the allotted a time a refuge.

As the skew angle of the intersection increased, pedestrian crashes tended to decrease. This could potentially be indicative that pedestrians are less inclined to cross at intersections with high skew or that pedestrians and drivers are more attentive at these locations.

Parking on the minor street was associated with increased crash risk. The presence of vehicles parked on the side of the road could obscure vision of pedestrians and motorists alike.

In the four-leg and joint models, the Bay region was shown to be associated with fewer pedestrian crashes, while the Grand region was shown to be associated with higher pedestrian crash frequency in each of the three models. These regional effects are possibly capturing unique characteristics about pedestrian crash reporting or pedestrian behavior in these areas.

3.5.2.2 Models for bicycle crashes by signalized intersection type

Table 31 illustrates the model results for bicyclist crashes at three-leg signalized intersections, Table 32 contains the model results for bicyclist crashes at four-leg intersections, while Table 33 contains a joint intersection model for bicyclist crashes where both three-leg and four-leg intersections were analyzed using the same model, with the two intersection types being differentiated using an indicator variable.

Table 31. Bicyclist crashes at three-leg signalized intersections

Parameter	Estimate	Std. Error	z-Value	p-Value
Intercept	-14.111	3.774	-3.740	<0.001
Natural log of bicycle commuter density	0.059	0.101	0.590	0.558
Natural log of major road traffic volume	0.912	0.377	2.420	0.016
Natural log of minor road traffic volume	0.178	0.115	1.550	0.120
Number of minor road driveways	0.228	0.090	2.530	0.012
Major leg one-way	-0.686	0.615	-1.120	0.265
University Region	1.254	0.372	3.370	0.001
Overdispersion Parameter	0.042	0.290		

Table 32. Bicyclist crashes at four-leg signalized intersections

Parameter	Estimate	Std. Error	z-Value	p-Value
Intercept	-7.680	1.866	-4.116	<0.001
Natural log of bicycle commuter density	0.174	0.068	2.546	0.011
Natural log of major road traffic volume	0.408	0.202	2.021	0.043
Natural log of minor road traffic volume	0.121	0.101	1.197	0.231
Census Tract Income ($10,000)	-0.100	0.074	-1.344	0.179
Number of minor road driveways	0.072	0.036	1.992	0.046
Major leg one-way	-0.719	0.358	-2.007	0.045
Grand Region	0.428	0.293	1.459	0.145
Bay Region	-0.705	0.400	-1.764	0.078
Southwest Region	0.373	0.302	1.232	0.218
University Region	0.580	0.283	2.049	0.041
Overdispersion Parameter	0.711	0.255		

Table 33. Joint model for bicyclist crashes at signalized intersections

Parameter	Estimate	Std. Error	z-Value	p-Value
Intercept	-10.276	1.967	-5.224	<0.001
Natural log of bicycle commuter density	0.150	0.059	2.539	0.011
Natural log of major road traffic volume	0.588	0.194	3.030	0.002
Natural log of minor road traffic volume	0.151	0.077	1.948	0.051
Census Tract Income ($10,000)	-0.055	0.058	-0.949	0.342
Three-leg intersection	-0.227	0.236	-0.963	0.335
Number of minor road driveways	0.099	0.034	2.938	0.003
Major leg one-way	-0.539	0.307	-1.759	0.079
Superior Region	0.478	0.418	1.142	0.253
North Region	0.527	0.344	1.531	0.126
Grand Region	0.613	0.311	1.974	0.048
Bay Region	-0.394	0.405	-0.972	0.331
Southwest Region	0.575	0.319	1.802	0.072
University Region	0.985	0.309	3.184	0.001
Overdispersion Parameter	0.647	0.219		

Cyclist crash frequency was shown to increase with the density of cyclist commuters in a census tract as well as the major and minor traffic volumes, which serve as exposure measures for the models. Similarly to the pedestrian models, cyclist crashes were shown to decrease as census tract income increased. From a geometric standpoint, three-leg intersections and major leg one-way streets were also shown to be associated with fewer crashes, a result that is also similar to that of the pedestrian models.

As the number of minor road driveways increased, so too did predicted crash frequency. This suggests that there is difficulty for motorists and cyclists in detecting conflicts at driveways.

From a regional stand point, the Bay region was associated with fewer crashes in both the four-leg and joint models, which is again similar to the pedestrian models. The University region was shown to be associated with increased crash frequency in all three cyclist models, while the Southwest and Grand regions were shown to be associated with increased crash frequency for the four-leg and joint models and the Superior and North were shown to be associated with increased crash frequency in the joint model. While regional differences are likely attributable to reporting behavior and activity level, it is worth noting that the North and Superior regions are the least populous in the state and that their respective regional indicators are reflective of the fact that for a specific volume, they are more dangerous than the metro Region.

3.4.2.3 Non-motorized Models for by signalized intersection type

Table 34 illustrates the model results for non-motorist crashes at three-leg signalized intersections, Table 35 contains the model results for non-motorist crashes at four-leg intersections, while Table 36 contains a joint intersection model for non-motorist crashes where both three-leg and four-leg intersections were analyzed using the same model, with the two intersection types being differentiated using an indicator variable.

Table 34. Non-motorized crashes at three-leg signalized intersections

Parameter	Estimate	Std. Error	z-Value	p-Value
Intercept	-14.298	2.905	-4.922	<0.001
Natural log of non-motorized commuter density	0.035	0.074	0.471	0.638
Natural log of major road traffic volume	1.057	0.290	3.646	<0.001
Natural log of minor road traffic volume	0.192	0.082	2.333	0.020
Census Tract median income ($10,000)	-0.148	0.076	-1.952	0.051
Major leg one-way	-0.591	0.461	-1.281	0.200
Major road median presence	-0.423	0.401	-1.056	0.291
North Region	0.436	0.456	0.956	0.339
Grand Region	0.741	0.377	1.965	0.049
Bay Region	1.032	0.352	2.929	0.003
University Region	1.338	0.355	3.769	<0.001
Overdispersion Parameter	0.043	0.206		

Table 35. Non-motorized crashes at four-leg signalized intersections

Parameter	Estimate	Std. Error	z-Value	p-Value
Intercept	-7.792	1.658	-4.701	<0.001
Natural log of non-motorized commuter density	0.160	0.058	2.756	0.006
Natural log of major road traffic volume	0.531	0.161	3.303	0.001
Natural log of minor road traffic volume	0.084	0.078	1.077	0.282
Census tract median income ($10,000s)	-0.061	0.065	-0.945	0.345
Number of minor road driveways	0.036	0.029	1.236	0.216
Major road median presence	-0.446	0.265	-1.681	0.093
Census tract median age	-0.014	0.014	-0.944	0.345
Minor road parking	0.344	0.197	1.747	0.081
Grand Region	0.638	0.232	2.751	0.006
Bay Region	-0.455	0.269	-1.693	0.090
University Region	0.383	0.208	1.841	0.066
Overdispersion Parameter	0.484	0.141		

Table 36. Non-motorized crashes at signalized intersections

Parameter	Estimate	Std. Error	z-Value	p-Value
Intercept	-7.958	1.440	-5.525	<0.001
Natural log of non-motorized commuter density	0.124	0.047	2.664	0.008
Natural log of major road traffic volume	0.563	0.138	4.084	<0.001
Natural log of minor road traffic volume	0.135	0.059	2.293	0.022
Three-leg intersection	-0.323	0.176	-1.831	0.067
Census tract median income ($10,000s)	-0.087	0.052	-1.677	0.094
Number of minor road driveways	0.053	0.027	1.981	0.048
Major leg one-way	-0.529	0.214	-2.470	0.014
Minor road median presence	-0.502	0.344	-1.462	0.144
Major road median presence	-0.381	0.225	-1.697	0.090
Skew angle	-0.008	0.005	-1.661	0.097
Census tract median age	-0.021	0.012	-1.702	0.089
Minor road parking	0.264	0.166	1.588	0.112
Grand Region	0.443	0.189	2.341	0.019
Bay Region	-0.587	0.243	-2.415	0.016
University Region	0.385	0.179	2.152	0.031
Overdispersion parameter	0.415	0.118		

The primary variable of interest in this analysis are the pseudo-exposure measures, pedestrian commuter density, bicyclist commuter density, and non-motorized commuter density.

Generally speaking, as this value increased, the number of combined pedestrian and cyclist crashes increased. Only one of the nine models estimated as a part of this study (three-leg signalized bike crashes) did not display a clear relationship between the commuter density and crash frequency. These measures certainly do not capture the exact behavior of non-motorized road users at any specific location to the same degree that directional counts would be able to, however, they do appear to be a reliable predictor of the general activity level in the area.

The traffic volume on the major and minor streets is shown to have a substantial effect on the frequency of non-motorized crashes at signalized intersections. This result should come as no surprise, given that higher volume intersections likely result in more turning movements which are a likely movement to result in a collision with a non-motorized road user.

Three-leg intersections tend to have a lesser frequency of crashes than their four-leg counterparts. Intuitively, this makes sense given that there are four areas that non-motorized users could potentially be using to navigate the intersection as opposed to three. In fact, the elasticity (or percent change) associated with this coefficient (for the joint intersection/non-motorized user model) is a 0.276, indicating that the percent difference in predicted crashes involving non-motorized road users between four-leg and three-leg intersections is only slightly greater than 25%, suggesting a nearly directly proportional relationship between the number of potential crossing areas to the number of crashes.

Geography was shown to play a part in non-motorized crash frequency. Perhaps unsurprisingly, intersections located in MDOT's University region were shown to be associated with elevated crash frequency. MDOT's Grand region was also shown to be associated with elevated non-motorized crash frequency which again is potentially attributable to a relatively active population in the area. Additionally, one factor that could be influencing the frequency of crashes involving non-motorized road users in these areas could be the reporting tendencies present there. Conversely, MDOT's Bay region was associated with lower frequencies of crashes involving pedestrians and cyclists. This may be attributable to general activity levels in the area or to differences in crash reporting practices.

Higher levels of average median income of the census tract was found to be associated with lower crash frequency. This result is consistent with the extant literature which suggests that people are somewhat less likely to walk or bike as their income increases. The results were relatively consistent across each of the nine models. There are several possible explanations as to why lower-income areas are more likely to experience crashes involving non-motorized users. First and foremost, it is possible that the infrastructure in these areas in not conducive to use by non-motorized users. Older signals may not have signal heads for pedestrian crossings or meet minimum crossing time criteria and sidewalks may be in disrepair.

Surprisingly, presence of bike lanes at intersections was not shown to have an effect on cyclist crashes. Generally speaking, bike lanes were not common in the data set during the study period. The minimal installation locations could have been further confounded due to selectivity bias, wherein locations with high levels of cyclists or cyclist crashes were chosen as locations for bike lane installation. Regardless, the impact of dedicated bike infrastructure could not be thoroughly investigated as a part of this study, however, it bears mentioning given the thrust of this research. The lack of bike infrastructure may be forcing more cyclists to the sidewalk, which helps to justify the estimation of joint non-motorized road user models.

Interestingly, the number of non-motorized road user involved crashes was shown to increase as the number of driveways on the minor leg of the intersection increased. There are a variety of potential explanations for this observation. First, increasing driveway counts on the minor leg of the intersection is indicative of higher levels of single family residential development. In these situations, people might be back out of their driveway or just be generally unaware of pedestrians or cyclists in the area. These types of behavior could lead to collisions where the non-motorized road user is struck where a sidewalk crosses a driveway. Additionally, these types of locations may be more prone to having visual obstructions, such as trees, shrubs, or fences which may make non-motorized users less conspicuous.

Intersections where the major leg has one-directional vehicular travel were less prone to crashes involving non-motorized road users. One-way roads are provide drivers fewer variables to consider when maneuvering their vehicle. Additionally, pedestrians do not have to consider vehicles travelling in both directions when crossing the major street, which also means reduced turning movements, helping to minimize the number of conflicts between vehicles and non-motorized road users.

The presence of medians on the major and minor roads was associated with lower crash frequency. By providing non-motorized road users with a refuge, only one direction of flow needs to be considered at a time during a crossing maneuver. Additionally, the presence of medians may also restrict the turning movements of vehicles into driveways immediately adjacent to an intersection.

The effect of skew angle on crashes involving non-motorized road users is somewhat counter-intuitive. As the degree of skew angle increases, crashes involving non-motorized road users tended to decrease. One potential explanation for this phenomenon is that with high-skew intersections, pedestrians and cyclists may tend to avoid crossing at these locations or that these locations.

The median age of people living within the census tract of an intersection was found to be inversely related to the expected non-motorized crash frequency of an intersection. There are a bevy of possible explanations for this finding. First, communities with older populations may be less likely to travel by non-motorized means. Second, younger median ages of census tracts are presumably due to the proliferation of young children within the community. Children may be specifically vulnerable to crashes whether on foot or on bicycle due to lack of experience as a road user and potential lack of attention paid to their surroundings.

Finally, the presence of parking on the minor street was found to be associated with elevated crash frequency. This is potentially due to the visual obstruction that parked cars create for both vehicles and non-motorized road users.

Beyond demonstrating the ability of the models to assess geometric and geographic characteristics associated with traffic crashes for pedestrians, cyclist, and combined non-motorists, this research examined the predictive ability of each of the models. In order to provide an apples-to-apples comparison, the models were all applied to the five-year aggregated data. In order to predict the number of pedestrian and cyclist crashes using the combined non-motorist models, the predicted number of non-motorist crashes was multiplied by the pedestrian proportion of the non-motorist density and cyclist proportion of the non-motorist density, respectively. In the same vein, the model developed using both intersection types with a binary indicator to designate the number of legs was applied to the three-leg and four-leg data sets separately. By utilizing these approaches, an apples-to-apples comparison of the various models was achieved. Table 37 presents the average predicted crashes, MAD, and MSPE for each of the pedestrian models developed in this paper, while Table 38 presents the cyclist models.

Table 37. Goodness of Fit Summary for Pedestrian Models

	Average Predicted Crashes		MAD		MSPE	
Model type	3SG	4SG	3SG	4SG	3SG	4SG
AADT Only Ped	0.050	0.086	0.275	0.448	0.374	0.686
Fully Specified Ped	0.052	0.085	0.260	0.430	0.357	0.654
Combined Intersection Ped	0.048	0.082	0.258	0.430	0.357	0.660
Separate Intersection Non-motorized Total Pedestrian Proportion	0.103	0.178	0.305	0.478	0.371	0.623
Combined Intersection Non-motorized Total Bicyclist proportion	0.051	0.064	0.287	0.556	0.411	1.158

Table 38. Goodness of Fit Summary for Cyclist Models

	Average Predicted Crashes		MAD		MSPE	
Model type	3SG	4SG	3SG	4SG	3SG	4SG
AADT Only Bike	0.027	0.089	0.282	0.561	0.419	1.112
Fully Specified Bike	0.054	0.108	0.279	0.547	0.378	1.058
Combined Intersection Bike	0.053	0.109	0.287	0.545	0.390	1.053
Separate Intersection Non-motorized Total Bicyclist Proportion	0.054	0.066	0.287	0.559	0.408	1.158
Combined Intersection Non-motorized Total Bicyclist proportion	0.051	0.064	0.287	0.556	0.411	1.158

Generally speaking, the models estimated by simultaneously considering both intersection types using an indicator variable performed similarly to the models that were estimated separately for each combination of crash and intersection type. Conversely, the models estimated for combined non-motorist crashes and then multiplying the predicted non-motorized crash total by the pedestrian proportion of non-motorized commuters and cyclist proportion of non-motorized commuters tended to perform worse.

3.6 Conclusions

The results of the analysis presented in this chapter represent a reasonable approach to predict pedestrian and cyclist crashes when pedestrian or cyclist volume data are unavailable. The methodology outlined in this paper serves to form a template with which road agencies can feasibly begin to evaluate intersections for pedestrian and cyclist safety. Separate SPFs were estimated for three crash categories: pedestrian crashes, cyclist crashes, and total non-motorized crashes. The models estimated for total non-motorized crashes were then multiplied by the proportion of the non-motorized commuters comprised of pedestrians and bicyclists so that model fit could be evaluated for each crash type and compared to the other models. Models for each crash type were estimated for three-leg signalized intersections, four-leg signalized intersections, and combined signalized intersections. Goodness of fit for the combined intersection models was evaluated separately for each intersection type.

In addition to demonstrating the usefulness of a publicly available data source to improve the crash prediction efforts for pedestrians and cyclists, this section serves to document how pedestrian and cyclist crashes are affected by similar infrastructure components, at least in urban and suburban intersections in Michigan. In addition to the volume components that are traditionally expected to have a large influence on the frequency of crashes, this research demonstrates that in general, median presence, one-way roads, increasing skew angle, and increasing census tract age are associated with decreased crash risk while parking and driveways are associated with lower crash risk. Additionally, various areas of the state that are more prone to non-motorized crashes were identified through the use of regional indicator variables.

CHAPTER 4. ACCESS POINT PROXIMITY TO CROSSROAD RAMP TERMINALS

4.1 Background

A principal concern pertaining to access management is controlling the location of driveways and intersections near the termination point of highway interchange off-ramps. Access point density is often identified as a primary contributor to poor safety performance on any type of corridor. Poor access management is associated with several undesirable impacts, including: increased frequency of crashes (including crashes with pedestrians and cyclists), and decreased roadway efficiency (Williams 2003). From the point of view of land owners and developers, locating intersections and driveways as close to ramp terminals seems ideal. Decades old design guidance allowed for access points as close as 100 feet (AASHTO 1991), however policy has changed over the past 26 years and many states have since adopted have adopted this range for use on their roadway networks, as demonstrated in Table 39 from NCHRP Report 420.

Table 39. Summary of Minimum Access Spacing Standards or Guidelines (adapted from Gluck, Levinson, and Stover 1999)

State/Province	Rural	Urban
1. Alabama	300 feet to access	100 feet to access
2. Alberta	1,400 feet from signal to access	Same
3. California	500 feet from ramp to access	Same
4. Illinois	410 feet minimum distance from ramp to nearest intersection	Same
5. Iowa	660 feet rural primary highway, 330 feet other road or street	170 feet urban
6. Kentucky	300 feet to access	100 feet to access
7. Maryland	Based on geometrics, speeds, volumes, presence of signals and queuing	Same
8. N. Dakota	AASHTO guidelines	AASHTO guidelines (100 feet)
9. Ohio	600 feet for diamond interchange, 1,000 feet for cloverleaf	
10. Oregon	200 feet from frontage road, 500 feet from ramp (suggested)	Same
11. Pennsylvania	AASHTO guidelines (300 feet)	AASHTO guidelines (100 feet)
12. South Carolina	500 feet desirable, 300 feet minimum	300 feet desirable, 150 feet minimum
13. Texas	AASHTO guidelines (300 feet)	AASHTO guidelines (100 feet)
14. Utah	300 feet to access	150 feet to access
15. Virginia	200 feet from entrance ramp	Same
16. West Virginia	300 feet to access	100 feet to access
17. Washington	300 feet to access	300 feet to access
18. Wisconsin	1,000 feet to access, (500 feet- minor roads)	500 feet to access
19. Wyoming	300 feet to access	150 feet to access

Increased spacing for access points that are located within close proximity to ramp terminals has been a necessity, as the extant literature clearly documents issues that arise on these corridors, such as complicated weaving movements and complex signal operations which can lead to increased crash frequency as well as operational congestion (Gluck, Levinson, and Stover 1999; Gluck and Lorenz 2010). It is easy to speculate that the aforementioned problems, which could be applied access spacing in general, may be exacerbated when drivers are transitioning from an uninterrupted flow roadway (such as a freeway) to a crossroad. Figure 23 provides insight as to the types of movements vehicles may have to accomplish prior to approaching the first access point on a multi-lane highway (Florida DOT 2014).

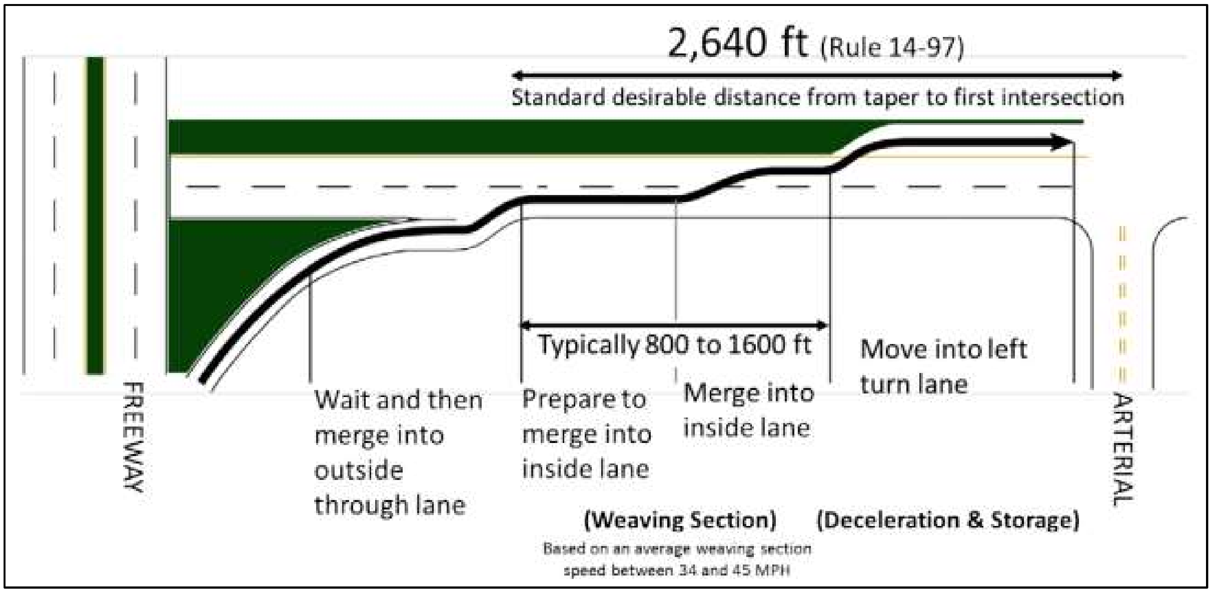


Figure 23. Distance between an off-ramp and first signalized intersection (Florida DOT 2014)

Currently, the Iowa DOT specifies a distance of 600 feet between ramp bifurcation points and first access points, however 300 feet is used in some instances depending on the land use near the interchange. These values are illustrated in Table 40.

Table 40. Iowa Design Guide Ramp to Access Point Spacing

English units

Roadway Type	Rural	Fringe	Built-up
Multi-lane Divided Highway	600’	600’	600’
Two-Lane Primary	600’	600’	300’
Secondary Road	600’	600’	600’
City Street	--	300’	300’

The purpose of this study is two-fold: to assess an appropriate distance at which access points can be located adjacent to ramp terminals, as well as to determine a minimum traffic volume threshold for the aforementioned distances to be applicable.

4.2 Literature Review

Various research studies have investigated issues related to the spacing of the first access point relative to the ramp intersection. A recurring theme in the extant literature is the effect of access point proximity on congestion. One research project aimed at developing guidelines for access point location in Florida through the use of computer simulations found that signals should be placed at minimum of ¼ mile from the ramp bifurcation point, while ½ mile should be used in cases where high development is expected (Washburn and Kondyli, 2006). A recent study utilized a mixed-integer non-linear model to develop non-traditional lane assignment through pavement markings and signal timings in an urban setting to minimize congestion at the off-ramp intersection (Zhao and Liu; 2016).

In addition to previous efforts that have focused on operational effects of access point location, the extant literature contains examples of studies focusing on the safety implications of access point proximity to ramp bifurcation points as well. A Virginia study utilized data from 186 access road sections to estimate negative binomial and least square linear regression models to examine the effect of access point spacing relative to the ramp bifurcation point (Rakha et al. 2008). The findings of this study were very pronounced: an increase in the spacing from 0 meters to 300 meters equated to an eight-fold decrease in crash rate, while an increase from 90 meters to 180 meters resulted in a 50 percent decrease in the crash rate. A subsequent report documented the similarity of the results between the negative binomial and linear regression approaches (Rakha et al. 2010).

Efforts have been made by researchers to combine the safety and operational effects of access point proximity near interchanges into monetary value. A study in Florida examined the relationship between distance to the first access point on a frontage road and the effect on congestion and crashes on the adjacent freeway (Williams et al, 2004). Key findings of this study included marked improvements in traffic flow when the distance to the first access point on a frontage road was 200 feet to 600 feet, general decreasing trend in crash frequency as the distance to the access point increased, and large benefit/cost ratios through the purchasing of land near interchanges ahead of development occurring.

4.3 Data Set Assembly

4.3.1 Interchange Manual Review and Identification of 'Ramp' Intersections

Interchanges were manually identified using an attribute query of the GIMS database to identify any segment considered to be a "ramp". This process resulted in the identification of all controlled-access highway to crossroad interchange locations, however, also included all system interchanges, including fully directional interchanges. Therefore, a preliminary manual review of all interchanges was conducted to identify such ramp terminal intersections. During the course of the manual review, other important information was collected including interchange type, traffic control on interchange ramps, whether roadways were divided and undivided, and whether or not a relevant spatial analysis could be conducted at each interchange. Spatial analyses cannot be conducted for fully directional interchanges (as shown in Figure 25) or for certain interchanges which were included in the database which are not freeway or expressway interchanges (examples are shown in Figure 26). Initially, it was found 406 interchanges in Iowa could be used as study locations for the spatial analysis. Subsequent quality control reduced the data set to

704 ramp corridors representing approximately 350 interchanges.

The manual review included collecting the following information:

- Distance to first driveway
- Distance to every intersection in study area (up to 1 mile from exit ramp bifurcation point)
- Distance to first field access
- Count of driveways to first and second intersections and total
- Median width at exit ramp bifurcation point, at first driveway, and at first intersection
- Median type at exit ramp bifurcation point, at first driveway, and at first intersection
- Volume of first access point
    - MSLINKs used to get volume from GIMS if first access point is an intersection
    - Volumes estimated based on driveway classification and the ITE Trip Generation Manual
- Turning movements of first driveway
- Side of the road for first driveway or intersection

Figure 24 shows the measurement of a driveway from the ramp bifurcation point. These distances were measured from the ramp point to the center of the driveway on the orthophoto. The measurement was recorded to the nearest foot.

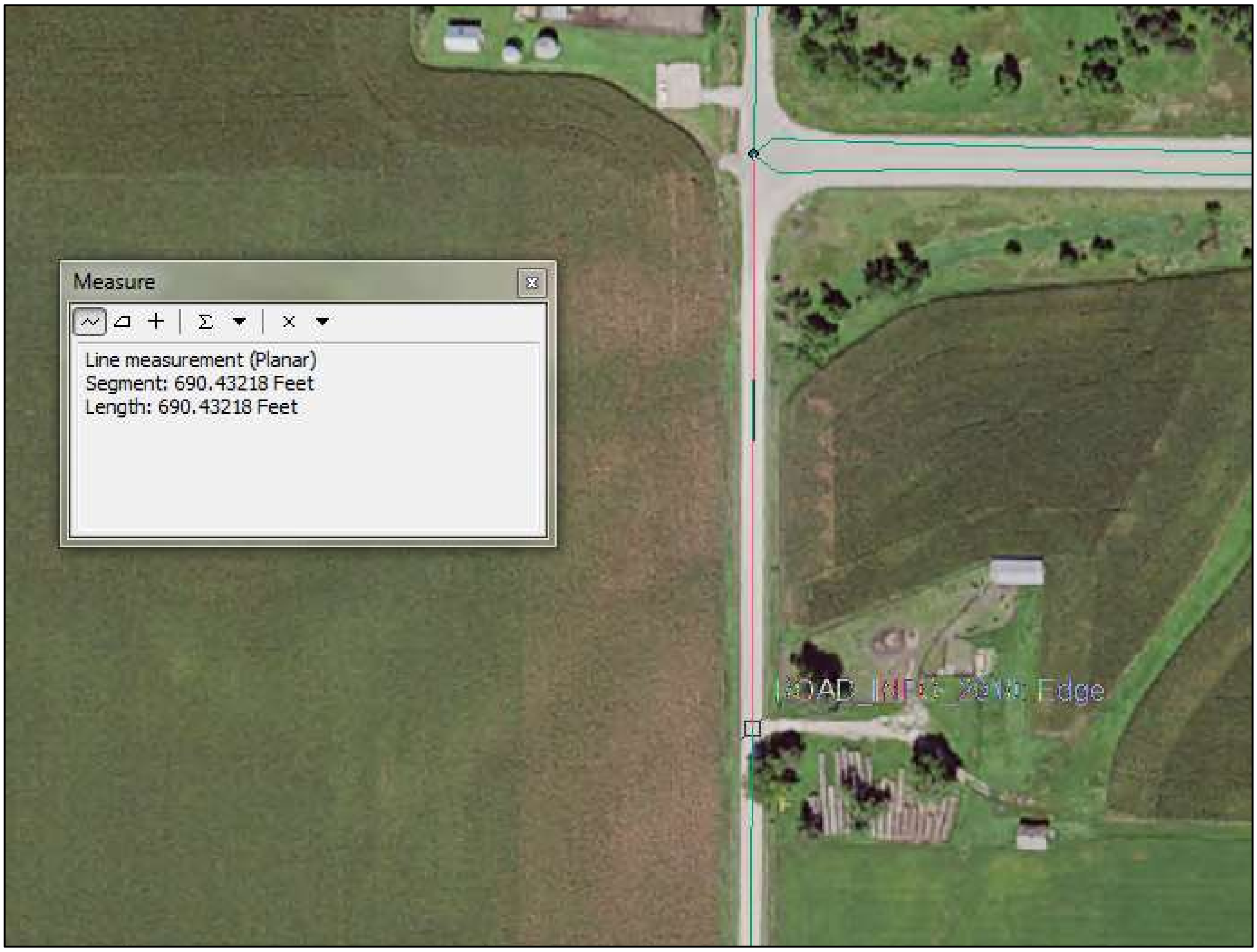


Figure 24. Example of GIS measurement to first driveway

Volumes were estimated for the driveway access points using the ITE Trip Generation Manual. For each classification determined by the researchers, the number of trips were calculated using the average values listed in the manual. Due to the difference in weekday and weekend traffic, the AADT for the classification was calculated using the average of the sum of the Saturday trips, Sunday trips, and 5 times the weekday trips. Table 41 contains the categories and AADT values used.

Table 41. Driveway AADT Estimates from ITE Trip Generation Manual

Classification	AADT	ITE Trip Generation Manual Page References
Apartments	1310	327
Camping	19	703,704
Cemetery	574	1097
Cell Tower	61	287,288
Church	251	1044
Construction	1124	99,104,106
Event Center	4390	853,855
Factory	4007	1509
Farm	1124	99,104,106
Farmhouse	10	290,295,297
Fast Food	1605	1821
Gas Station	1348	1888
Golf	739	754,759,761
Hospital	4234	1152
Hotel	2054	571,576,578
Maintenance	1124	99,104,106
Mechanic	41	1880
Office Building	2341	1196
Park	714	660,663,668,670
Parking	1152	82
Restaurants	1031	1794
School	1128	933
Shopping Center	14855	1497
Single House	10	290,295,297
Single Restaurant	843	1794
Storage	1382	208,213,215
Store	2052	1693
Truck Stop	1053	1896, 1794
Vehicle Dealership	936	1519

As these volumes are representative of the average example of the category, in some cases, particularly in the more rural areas of the state, the AADT estimated for the driveway exceeded that of the road the driveway was located on. In these instances, if the driveway was not opposite another roadway, the volume on the driveway was constrained to be equal to the volume on the main road.

Throughout the process of assembling and reviewing the data set, several types of interchanges were identified which had to be excluded from the analysis. Fully directional interchanges needed to be excluded as the lack the requisite crossroad on which to have an access point. An example of a fully directional interchange is shown in Figure 25.

Figure 25. Example of a Fully Directional Interchange (I-80 and I-29)

The attribute query utilized to identify interchanges also identified non-interchange ramps, such as those shown in Figure 26. As the scope of this study was focusing on interchanges from uninterrupted flow facilities to crossroads, these sites were excluded from the analysis as well.

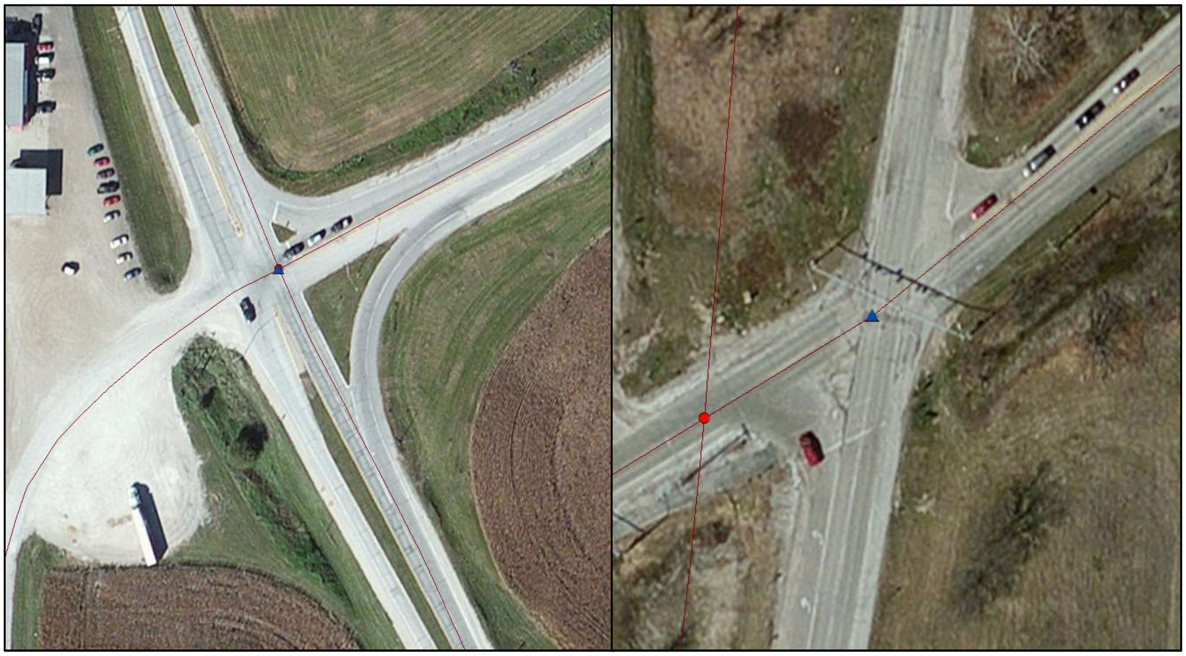


Figure 26. Examples of Non-Relevant 'Interchanges' Found in Interchange Database

4.3.2 Linear Referencing System

The Iowa DOT GIMS database contains information for individual GIMS segments, commonly referred as MSLINKs, which is the name given to their unique identification field. This level of data is not a suitable analysis unit to investigate the effect of access point proximity relative to the ramp terminal, which necessitated the development of a unique procedure of combining the data into analyzable format. This was accomplished through the development of a linear referencing system, a system in which roadway features, traffic crashes, etc. are all located on along a route corresponding to the centerline of the roadway. The beginning of the route is assigned a mile point of 0.0, while the end of the route is assigned a mile point equal to the route's length. Any point along the route can then be located based on the distance along the centerline from the beginning mile point. This approach is necessary to account for curved roadways.

Unique routes were created for each off-ramp terminal intersection. The route began (or ended depending on route directionality) at the point where the crossroad passed over or under the centerline of the controlled-access facility. The GIMS database does not have fields to facilitate linear referencing to the extent required for this analysis, therefore, all relevant GIMS segments (segments which represent the crossroad) were manually identified. This procedure required a large effort on the part of the research team, as any individual route in this analysis is potentially comprised of many separate GIMS segments. Figure 27 illustrates a typical route southbound at a diamond interchange, which is comprised of five segments of varying length.

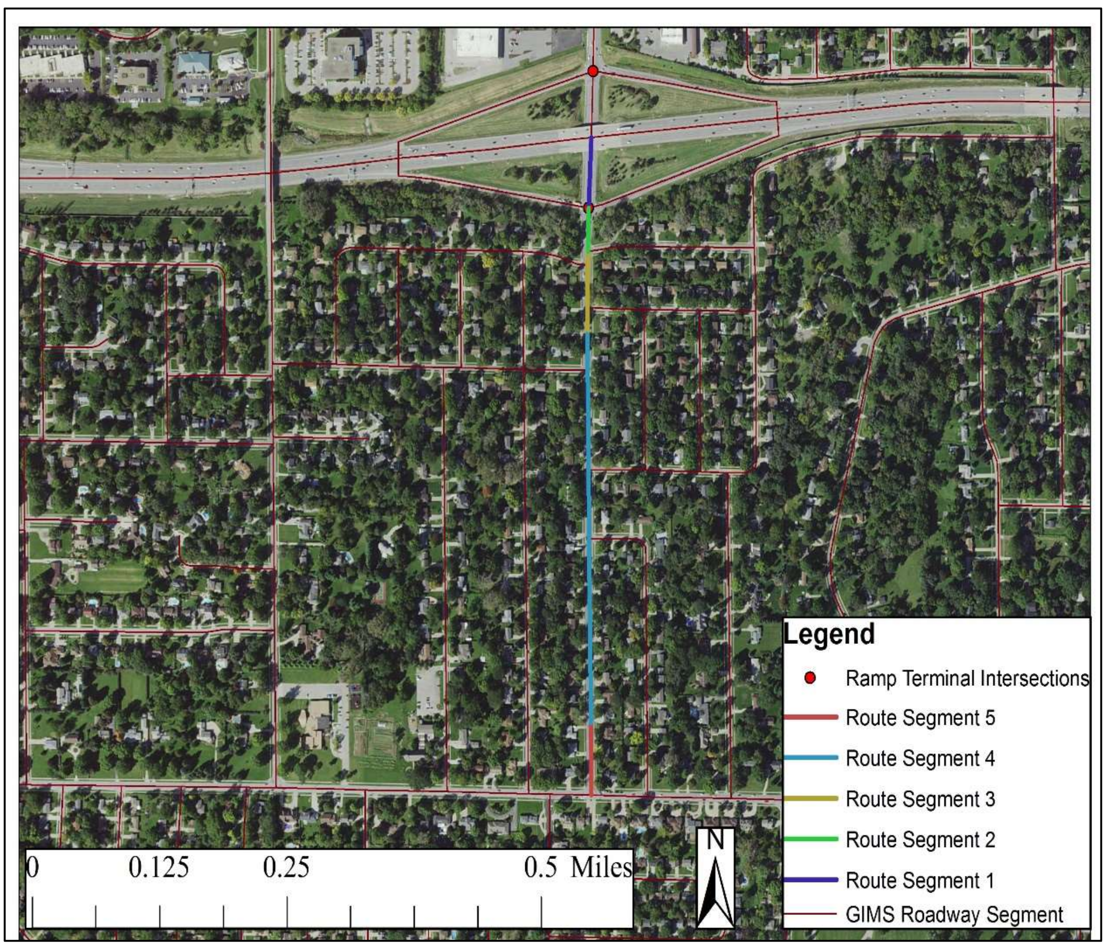


Figure 27. GIMS segments comprising a route

The individual GIMS segments are then dissolved in ArcGIS into one continuous segment. Two new fields were then added to the dissolved segments: a beginning point and an end point. The beginning point was set to equal zero for all of the segments while the end point was set to equal the calculated distance of the dissolved segments (in feet). The Create Routes tool in ArcGIS was then utilized to convert the dissolved segments into routes using the "TWO_FIELDS" option, with the beginning point serving as the "From-Measure Field" and the calculated end point serving as the "To-Measure Field", as shown in Figure 28. An example route with the ramp terminal intersections, crashes, and distance points for reference this analysis is shown in Figure 29.

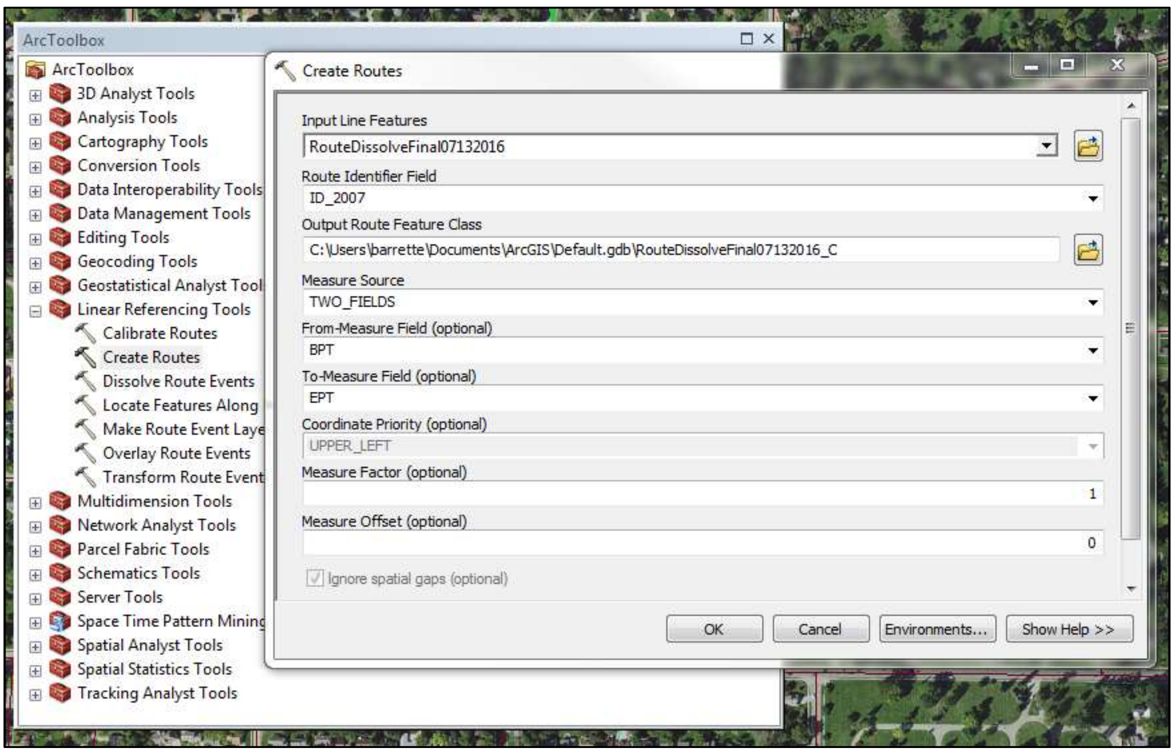


Figure 28. Screen shot of the Create Routes Tool in ArcGIS

Figure 29. Typical Crossroad Route

Following the creation of the routes, the position of relevant features (e.g., crashes, ramp terminal intersections, and GIMS segments) was identified by using the "Locate Features Along Routes" operation in ArcGIS as shown in Figure 30.

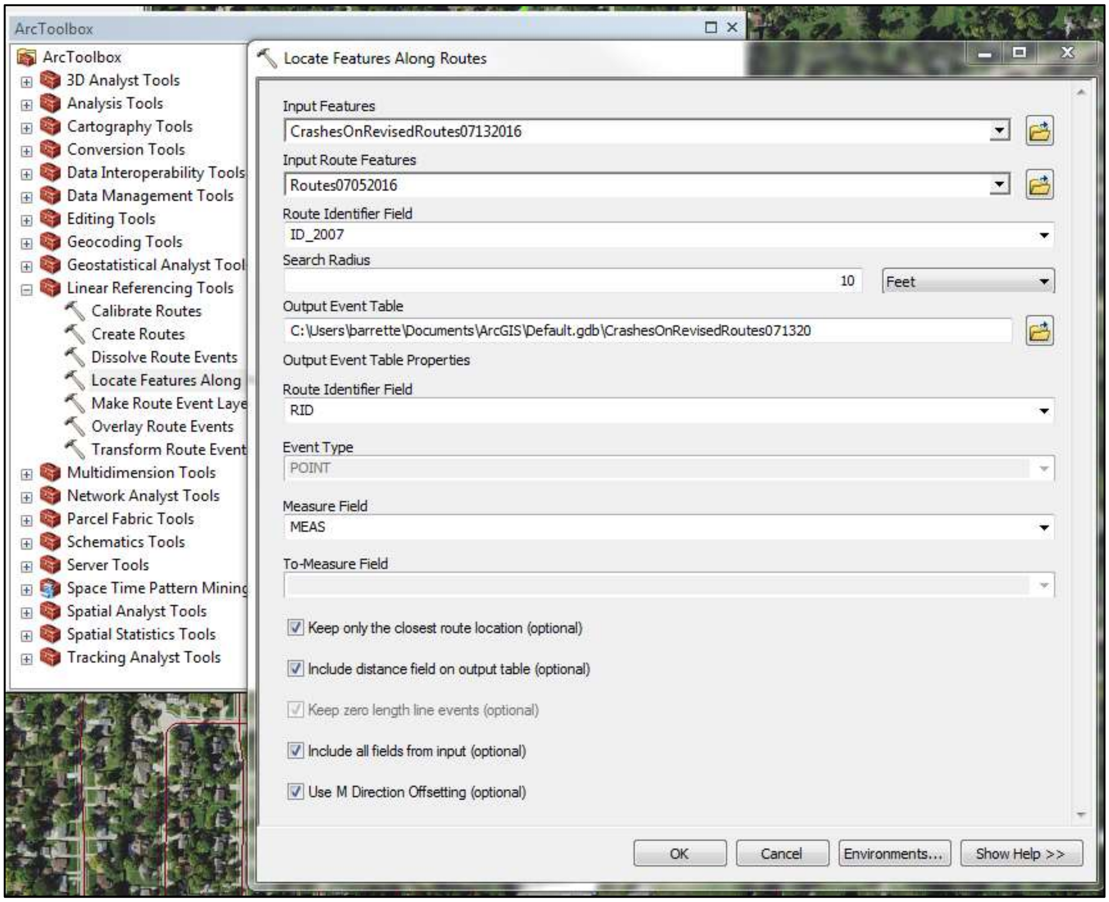
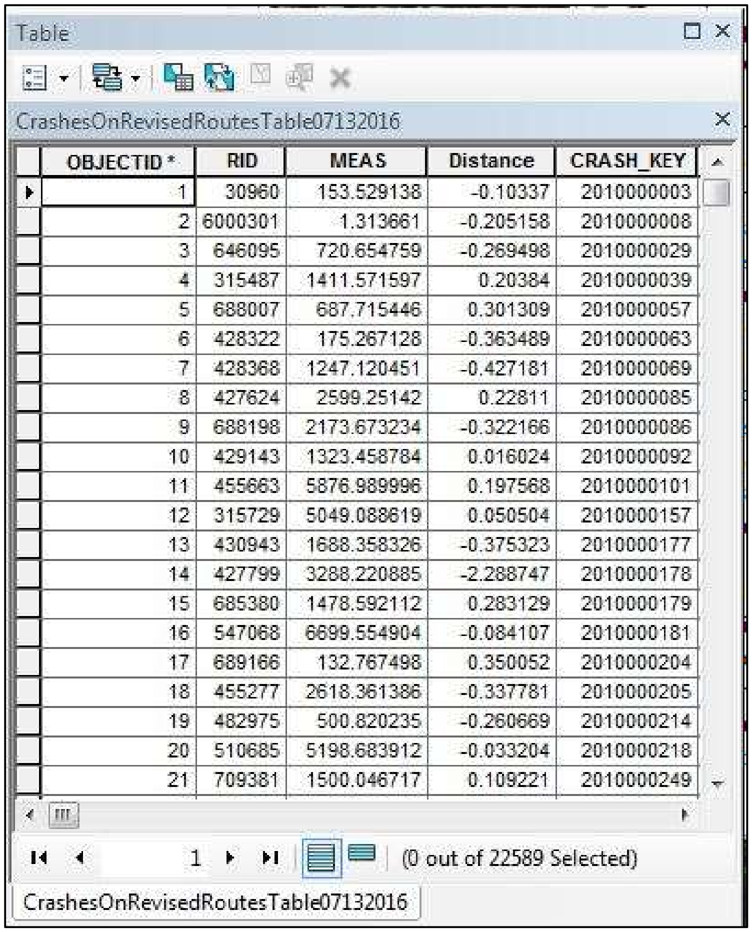


Figure 30. Locate Features Along Routes ArcGIS Interface

The result of the Locate Features Along Routes procedure was the creation of several ArcGIS tables, such as the one shown in Figure 31.



OBJECTID *	RID	MEAS	Distance	CRASH_KEY
1	30960	153.529138	-0.10337	2010000003
2	6000301	1.313661	-0.205158	2010000008
3	646095	720.654759	-0.269498	2010000029
4	315487	1411.571597	0.20384	2010000039
5	688007	687.715446	0.301309	2010000057
6	428322	175.267128	-0.363489	2010000063
7	428368	1247.120451	-0.427181	2010000069
8	427624	2599.25142	0.22811	2010000085
9	688198	2173.673234	-0.322166	2010000086
10	429143	1323.458784	0.016024	2010000092
11	455663	5876.989996	0.197568	2010000101
12	315729	5049.088619	0.050504	2010000157
13	430943	1688.358326	-0.375323	2010000177
14	427799	3288.220885	-2.288747	2010000178
15	685380	1478.592112	0.283129	2010000179
16	547068	6699.554904	-0.084107	2010000181
17	689166	132.767498	0.350052	2010000204
18	455277	2618.361386	-0.337781	2010000205
19	482975	500.820235	-0.260669	2010000214
20	510685	5198.683912	-0.033204	2010000218
21	709381	1500.046717	0.109221	2010000249

Figure 31. Crashes Located Along Routes

There are three fields that are created through the Locate Features Along Routes operation: RID, MEAS, and Distance. RID is a unique route identification number, MEAS is the location of a feature along a route, and Distance specifies how far off of the route a feature is actually located. This process was necessary so that the distance between roadway features, specifically between traffic crashes and the ramp terminal intersection, would be measured along the centerline of the roadway as opposed to a straight-line distance, which was crucial in situations where the roadway curved prior to the first access point. The linear referencing system allowed for the determination of what proportion of the distance between the ramp terminal intersection and the first access point each GIMS segment accounted for. Continuous properties of each GIMS segment, such as traffic volume, were then calculated as a weighted average for the route, while categorical properties for each segment were applied to each route as a length-weighted mode. The use of the length-weighted mode was designed to prevent instances of a few short segments being emphasized more than one long segment.

4.3.3 Data Set Creation

The resulting data set, referred to as the First Access Point (FAP) set due to simultaneous efforts to create other data sets that are beyond the scope of this paper, contains information about the roadway up to the first access point as well as more detailed information about that first access point, including estimated volumes, classification, turning movements, and distance from ramp bifurcation point. This data set was used to examine the relationship between crashes and distance to the first access point independent of anything that occurred after that access point. The five years of data (2010-2014) collected for 702 corridors represent a total sample size of 3,516 observations.

Table 42. Descriptive Stats for FAP Data Set (n=3,516)

	Average	Std. Dev.	Minimum	Maximum
First Access Point Distance	760.129	638.498	1.000	4329.100
Traffic Volume	5936.359	6924.324	10.000	44900.000
Truck Route	0.322	0.554	0.000	2.000
Average Median Width	5.159	12.634	0.000	138.316
Average Number of Lanes	2.853	1.167	2.000	9.000
Divided Proportion	0.241	0.412	0.000	1.000
Average Surface Width	35.770	15.648	16.000	119.912
Average Speed Limit	47.765	9.822	24.994	62.168
Outside Shoulder Width	5.969	3.822	0.000	16.000
Inside Shoulder Width	0.596	1.847	0.000	12.000
FA-Driveway	0.485	0.500	0.000	1.000
FA-Signalized	0.115	0.319	0.000	1.000
FA-Three Leg	0.263	0.440	0.000	1.000
FA-Four Leg	0.250	0.433	0.000	1.000
FA-Five Leg	0.001	0.038	0.000	1.000
No Left Out	0.027	0.162	0.000	1.000
Yield Controlled	0.011	0.106	0.000	1.000
Commercial DW	0.226	0.418	0.000	1.000
Full Movement	0.972	0.166	0.000	1.000
4 Leg Access Point	0.512	0.500	0.000	1.000
3 Leg Access Point on Left	0.127	0.333	0.000	1.000
3 Leg Terminal Access Point	0.017	0.130	0.000	1.000
Three Leg Right Access Point	0.394	0.489	0.000	1.000
Off Ramp AADT	1783.176	2367.084	10.000	16700.000
On Ramp AADT	1661.417	2428.329	0.000	29400.000
Average Ramp AADT	1722.297	2099.221	35.000	18200.000
Access Point AADT	1093.728	2371.148	0.000	20300.000
Right In Right Out	0.026	0.158	0.000	1.000
Left Out Only	0.001	0.038	0.000	1.000
Parallel Parking	0.083	0.276	0.000	1.000
Driveway Through	0.026	0.158	0.000	1.000
Signalized Access Point	0.115	0.319	0.000	1.000
All Way Stop Access Point	0.007	0.084	0.000	1.000
Yield/Uncontrolled Access Point	0.001	0.038	0.000	1.000
Signalized Ramp	0.216	0.412	0.000	1.000
Yield Controlled Ramp	0.038	0.192	0.000	1.000
Total Crashes	1.227	2.740	0.000	24.000

4.4 Statistical Analysis

4.4.1 Visual Analysis

To address access spacing relative to the location of freeway ramp bifurcation points, crash rates were calculated for all crashes occurring between the ramp bifurcation point and the first access point travelling away from the freeway along the crossroad. Figure 32 displays the crash rates between the ramp bifurcation point and first access point plotted against the distance to the first access point. This plot appears to indicate an inflection point near 0.1 miles, indicating that locations where access points are located at less than this distance experience elevated crash rates. This can potentially be attributed to conflicting movements being located closely together as well as driver expectations and behavior when leaving a limited access facility.

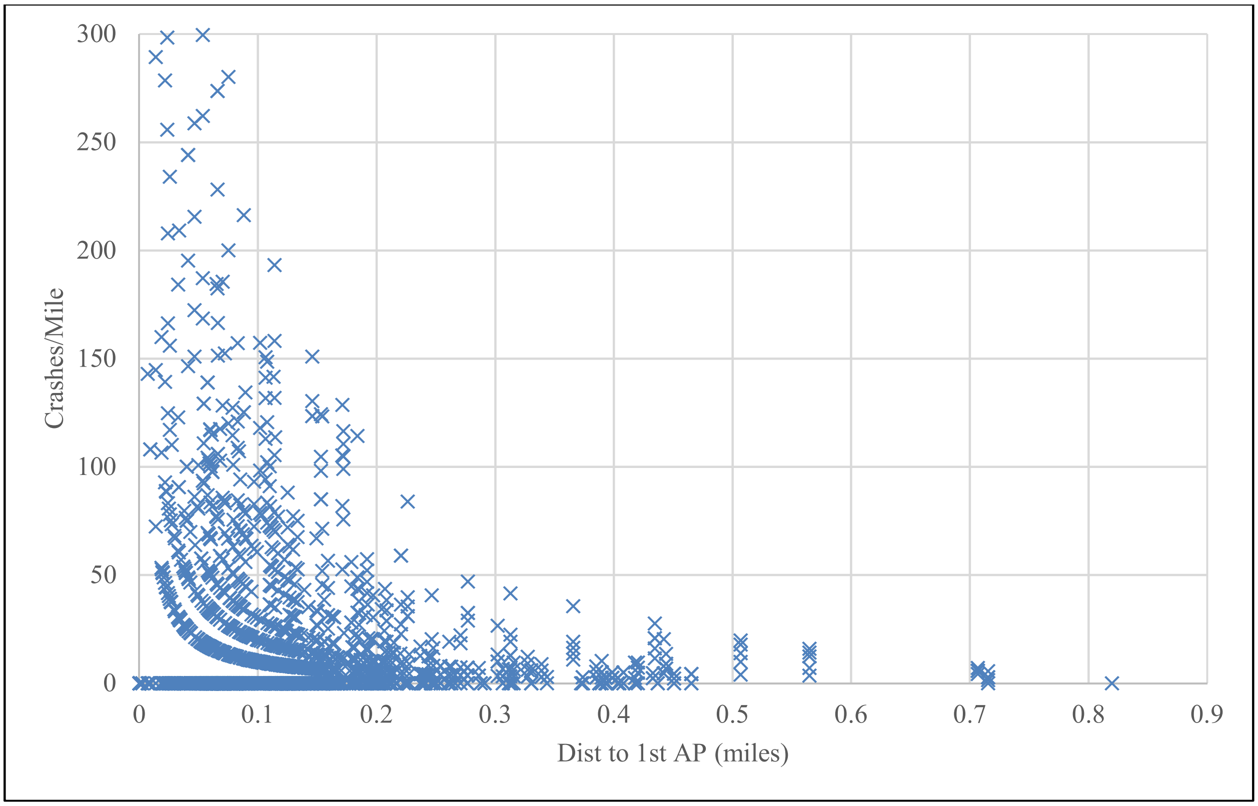


Figure 32. Crashes per Mile versus Distance to First Access Point at Ramp Bifurcation Points

Examining the crashes per mile versus distance to first access point within various AADT ranges provides some insight into the interaction between crossroad volume and access spacing. Figure 33 presents the same information as Figure 32, however the data points are limited to those site years with an AADT between 1000 vehicles per day and 1500 vehicles per day.

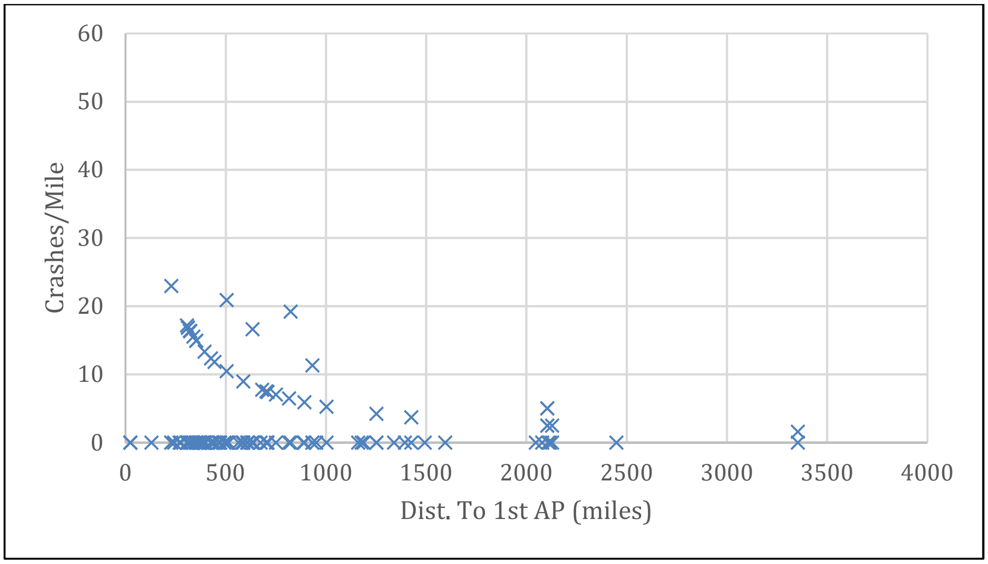


Figure 33. Crashes per Mile versus Distance to First Access Point for 1000 to 1500 Vehicles per Day

As can be seen in Figure 33, bands clearly begin to form that are generally associated with specific traffic volume ranges. Figure 34 presents the similar data to Figure 32 and Figure 33, with the traffic volume restricted to between 5000 and 6000 vehicles per day.

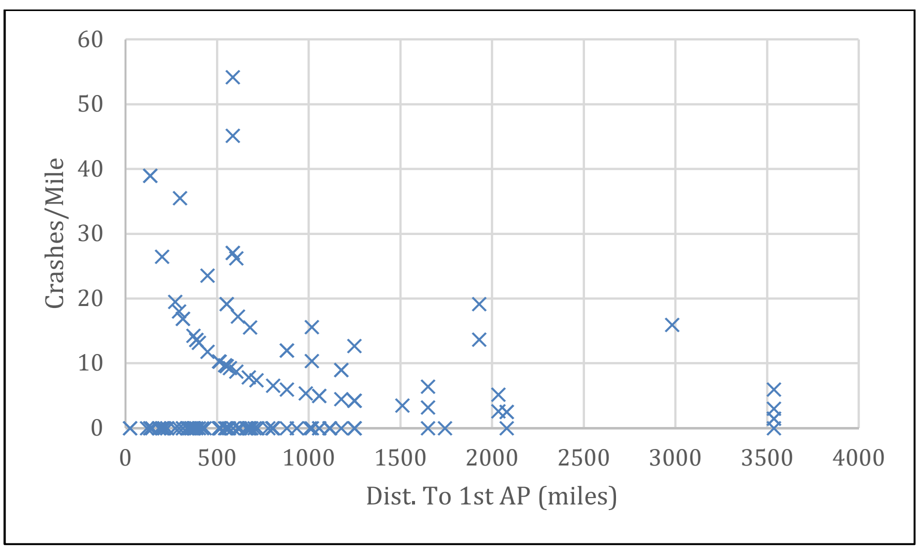


Figure 34. Crashes per Mile versus Distance to First Access Point for 5000 to 6000 Vehicles per Day

In Figure 34, there are two distinct curved bands, with a 3rd beginning to form as well. Generally speaking, as the traffic volume presented in the figures increases, the bands appear at higher crash rate magnitudes. Collectively, these plots demonstrate the relationship between traffic volume, access point proximity, and crash rate.

4.4.2 SPF Development

To discern the effects of the distance an access point is from the ramp bifurcation point, two models were estimated to determine how the crash risk changes as the distance to the first access point changes. The first model is focused on developing guidelines regarding access point spacing through the use of a series of binary indicators representing the location of the first access point in 100 feet increments. The second model was estimated using the distance as a continuous variable, along with the volume of traffic on the crossroad and the volume of traffic for the access point.

These SPFs take the form of generalized linear models. As crash data are comprised of non-negative integers, traditional regression techniques (e.g., ordinary least-squares) are generally not appropriate. Given the nature of such data, the Poisson distribution has been shown to provide a better fit and has been used widely to model crash frequency data. In the Poisson model, the probability of intersection i experiencing y_i crashes during a one-year period is given by Equation 22,

$$P(y_i) = \frac{EXP(-\lambda_i)\lambda_i^{y_i}}{y_i!} \tag{22}$$

where $P(y_i)$ is probability of intersection i experiencing y_i crashes and λ_i is the Poisson parameter for intersection i, which is equal to the segment's expected number of crashes per year, $E[y_i]$. Poisson models are estimated by specifying the Poisson parameter λ_i (the expected number of crashes per period) as a function of explanatory variables, the most common functional form being given by Equation 23,

$$\lambda_i = exp(\beta X_i) \tag{23}$$

where X_i is a vector of explanatory variables and β is a vector of estimable parameters.

A limitation of this model is the underlying assumption of the Poisson distribution that the variance is equal to the mean. As such, the model cannot handle overdispersion wherein the variance is greater than the mean. Overdispersion is common in crash data and may be caused by data clustering, unaccounted temporal correlation, model misspecification, or ultimately by the nature of the crash data, which are the product of Bernoulli trials with unequal probability of events (Lord 2006). Overdispersion is generally accommodated through the use of negative binomial models (also referred to as Poisson-gamma models).

The negative binomial model is derived by rewriting the Poisson parameter for each intersection as shown in Equation 24,

$$\lambda_i = exp(\beta X_i + \varepsilon_i) \tag{24}$$

where *EXP* (ε_i) is a gamma-distributed error term with mean 1 and variance α. The addition of this term allows the variance to differ from the mean as shown in Equation 25:

$$VAR[y_i] = E[y_i] + \alpha E[y_i]^2 \tag{25}$$

The negative binomial model is preferred over the Poisson model since the latter cannot handle overdispersion and, as such, may lead to biased parameter estimates (Lord and Park 2008). Consequently, the *HSM* recommends using the negative binomial model for the development of SPFs.

If the overdispersion parameter (α) is equal to zero, the negative binomial reduces to the Poisson model. Estimation of λ_i can be conducted through standard maximum likelihood procedures. While alternatives, such as the Conway-Maxwell model, have the advantage of accommodating both overdispersion and underdispersion (where the variance is less than the mean) (Lord and Mannering 2010), the negative binomial model remains the standard in SPF development.

Due to the presence of repeated observations resulting in temporal correlation among observations, random-effect models were estimated. Random effects are considered by re-writing the constant term in Equation 26 as follows:

$$\beta_{0i} = X\beta_0 + \omega_i \tag{26}$$

where ω_i is a randomly distributed random effect for intersection *j* and all other variables are as previously defined.

In contrast to intersection SPFs, where the measures of exposure are vehicle and/or non-motorized road user volume, length is typically included, commonly as an offset variable (coefficient restricted to equal 1) as an additional measure of exposure in models for roadway segments. The principal characteristic of interest in this study was the segment length (the distance between the ramp bifurcation point and the first access point), therefore a unique functional for SPFs was utilized and is shown in the following equation:

$$\lambda = exp(1 * ln(Length) + \beta_{Leng} \quad * ln(Length) + \beta_{CRAADT} * ln(CRAADT) + \beta_{APAADT} * ln(APAADT) + \boldsymbol{\beta X}) \tag{27}$$

where *Length* is the distance between the ramp terminal intersection and the nearest access point, β_{Length} is the coefficient associated with *Length*, *CRAADT* is the traffic volume of the crossroad running perpendicular to the freeway, β_{CRAADT} is the coefficient associated with the crossroad volume, *APAADT* is the volume of traffic entering the roadway from the access point, β_{APAADT} is the coefficient associated with access point volume, $\boldsymbol{\beta}$ is the vector other model coefficients, and $\boldsymbol{X}$ is the vector of the remaining explanatory parameters.

4.5 Results and Discussion

In order to develop useful guidance for road agencies, two crash prediction models were estimated. In an effort complimentary to the series of crash prediction model previously described, a detailed crash prediction model was developed treating the distance to the first access point as a continuous variable. While the interval based model is beneficial for identifying distance thresholds, the continuous model captures the true relationship of the data (e.g., there is not necessarily a huge difference between distances of 399 feet and 401 feet). Data was collected on an individual year basis for this analysis, therefore to account for correlation between repeated observations at individual sites, a random-intercept negative binomial model was estimated. The results of the model estimation are given in Table 43.

Table 43. Crash prediction model with distance as a continuous variable

Parameter	Estimate	Std. Error	z-Value	p-Value
Intercept	-10.049	0.484	-20.770	0.000
Natural log of crossroad AADT	0.786	0.052	15.120	0.000
Natural log of 1st access point AADT	0.082	0.024	3.410	0.001
Natural log of distance to first access point	-0.659	0.038	-17.250	0.000
First access point on right side	-0.335	0.096	-3.500	0.000
First access point is terminal intersection	-0.654	0.355	-1.850	0.065
Signalized first access point	0.374	0.131	2.860	0.004
Signalized ramp	0.833	0.119	6.980	0.000
Yield controlled ramp	0.384	0.194	1.970	0.048
Overdispersion parameter	0.049	0.015		

The variance associated with the random intercept was estimated to be 0.5975. Additionally, the log-likelihood associated with the random intercept model was determined to be -3515.15. In comparison to the log-likelihood of the same model estimated using naïve-pooled negative binomial framework (-3778.05), the random intercept model provides superior fit. An analysis of deviance test was subsequently performed, which confirmed that the random-intercept model is a significantly better fit than the naïve-pooled model. This information suggests that while a wide variety of parameters was found to significantly affect the crash rate at these sites (and are discussed in the following paragraphs), there are still unobserved site-specific parameters that also have an impact on crash rate.

The model estimated for this study examined the volume of the crossroad as well as the volume at the first access point. As one would expect, both of these volumes were found to have a substantial impact on crash frequency along the corridor. It is interesting to note that the parameter associated with the first access point traffic volume is relatively small. This suggests that the effect of access point volume increases rapidly at low volumes and then plateaus as access point volume increases. This is noteworthy because it implies that access points present crash risk to the corridor with relatively little regard for the amount of vehicles that utilize them.

The primary variable of interest in this study was the distance to the first access point from the point of ramp bifurcation. This site characteristic was included in the model as an offset term, as well as a separate estimable parameter. The offset term is constrained to a value of one, effectively resulting in the estimation of crash rates on a per mile basis. The estimated parameter associated with this characteristic was negative, which indicates that by increasing the distance to the first access point, the expect crash rate between the ramp bifurcation point and first access point decreases.

A variety of access point configurations were examined to ascertain best practices with regards to movement restrictions. Two approach configurations, driveways and three-leg intersections on the right side (right turn for vehicles traveling away from the highway on the crossroad) as well as three-leg terminal intersections (the crossroad ends in a t-intersection) were shown to have lower crash rates in comparison to the base scenario of the first access point being a four-leg intersection. Movement restrictions, such as right-in, right-out were also considered in the model estimation, but were not found to have a significant effect on crash rate. This is potentially due to this data set focusing explicitly on the first access point, while full corridor investigations could find that roadways with higher densities of restricted movement access points perform better than corridors with full-movement access points.

Finally, traffic control was shown to play a substantial role in crash rate. Signalized ramps and first access points were both associated with elevated crash rate. This is fairly intuitive, as these scenarios are most commonly present at locations with high traffic volumes, and therefore, the highest crash risk. Somewhat counterintuitively, yield controlled ramps (slip ramps) were found to be associated with higher crashes as well. This again is likely due to these being in locations where a large volume of vehicles are exiting the highway.

There were two principal foci behind this investigation that can be addressed graphically using the estimated crash prediction model: 1) illustrate the relationship between the spacing between the ramp bifurcation point and the first access point, and 2) illustrate the role that crossroad traffic volume plays on the rate or crashes at a given location. Figure 35 is the graphical interpretation of the crash prediction model with crossroad volume plotted on the x-axis.

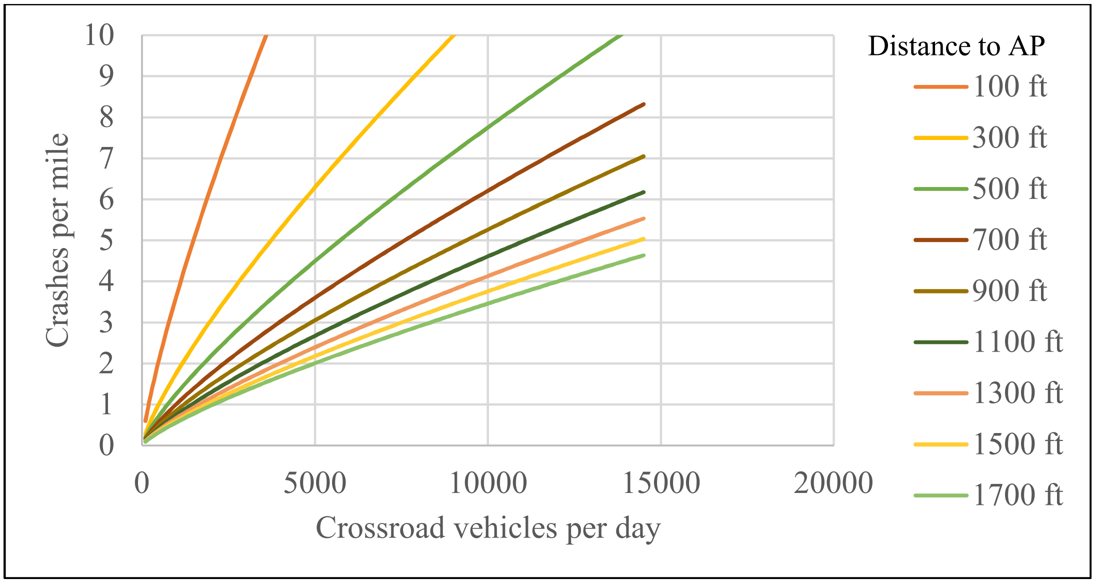


Figure 35. Crashes per mile versus crossroad traffic volume by distance to access point, access point volume of 100 vehicles per day

Figure 35 demonstrates that as traffic volume on the crossroad increases, crash rate also increases. This figure again demonstrates that the rate of traffic crashes increases much more rapidly with respect to traffic volume when the first access point is near to the ramp bifurcation point. The differences in crash frequency attributable to the distance between the ramp terminal intersection and first access point beyond 1,100 feet are not very pronounced.

Figure 36 is a zoomed version of Figure 35 emphasizing the low end of the volume spectrum and provides critical insight into a low-volume threshold at which was not able to be addressed with the previous plots.

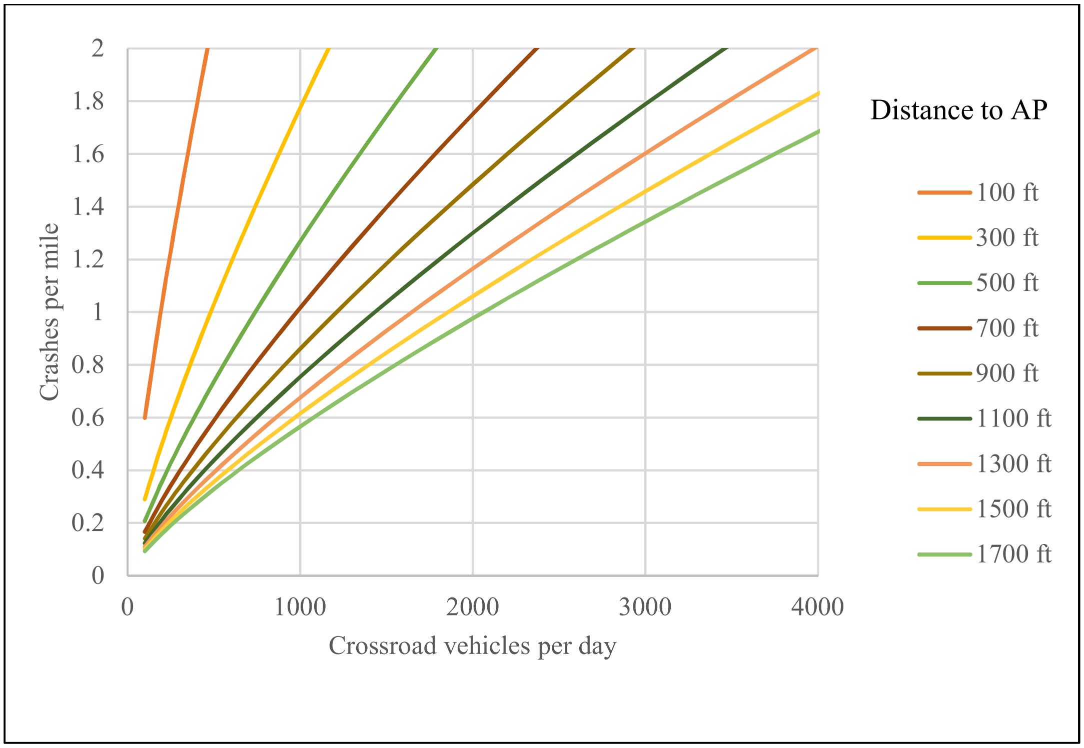


Figure 35. Low crossroad volumes, crashes versus crossroad volume

Figure 36 illustrates that from a practical sense, there is still a marked difference in crash rate between a corridor with crossroad volumes of 100 vehicles per day when the access point is located at 100 feet versus 300 feet (0.3 crashes per mile). The difference in crash rate when the distance between the ramp terminal intersection and first access point is increased from 300 feet to 500 feet (0.08 crashes per mile) is not nearly as pronounced. This figure further illustrates that the crash rate on a roadway is more sensitive to traffic volume when the distance between the ramp terminal and first access point is smaller, suggesting that high volume access points should generally be located further from the ramp terminal. Figure 37 again presents the SPF graphically, however, this the distance to the first access point is plotted on the X-axis.

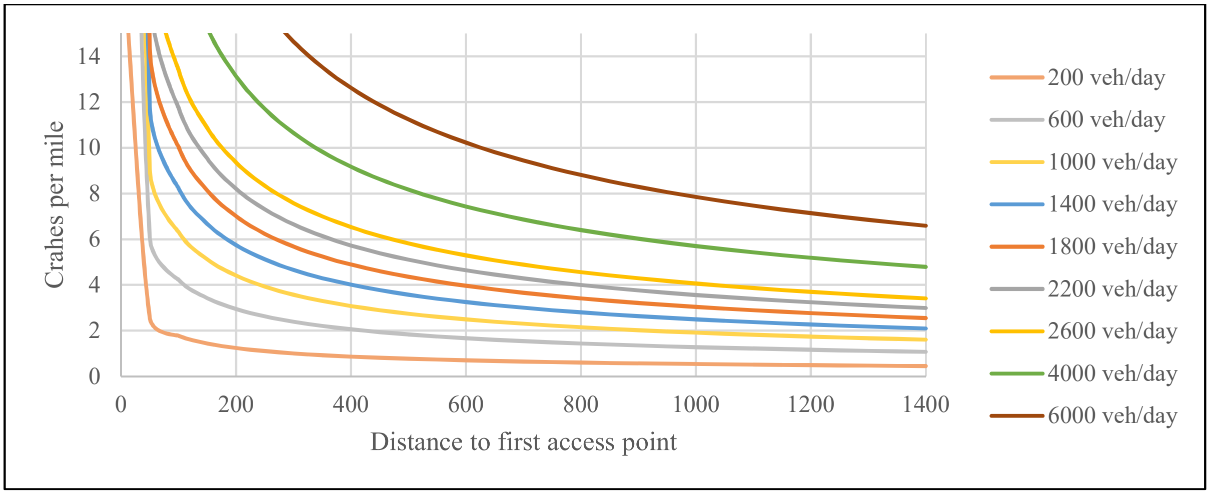


Figure 36. Crashes per mile versus distance to the first access point for various cross road volumes

Figure 37 is based on a right-side access point with a volume of 100 vehicles per day. Each bar represents a specific cross road traffic volume. This plot is useful for identifying inflection points in the curve which represent a substantial change in safety performance. For instance, the lowest curve (200 crossroad vehicles per day) is relatively unaffected by increasing the distance to the access point beyond 100 feet. As crossroad traffic volume increases, a wider range of access point distances have a pronounced effect on crash rate. Collectively, these plots and figures provide insight into the relationship between crash rate, the distance between the ramp intersection and the first access point, and the traffic volume present at the first access point.

The second model was developed on aggregated crash data for the five years of the study period. This model utilized a series of indicator variables for the distance between the ramp bifurcation point and the first access point. These ramp distance indicators were estimated as interaction terms with access point volume, as intuitively, one might expect that the effect of access point volume is highly related to the distance to the access point, i.e., locating a high-volume access point close to the ramp bifurcation point would have a much more pronounced effect than locating a low-volume access point at the same location. Results of the model are shown in Table 42.

Table 44. Five-year aggregated SPF results

Parameter	Estimate	Std. Error	z-Value	p-Value
Intercept	-14.205	0.413	-34.418	<0.001
Natural log of crossroad AADT	0.801	0.052	15.476	<0.001
First access point on right side	-0.347	0.092	-3.757	<0.001
First access point is terminal intersection	-0.626	0.343	-1.825	0.068
Signalized first access point	0.357	0.123	2.912	0.004
Signalized ramp	0.793	0.116	6.835	<0.001
Yield controlled ramp	0.340	0.188	1.805	0.071
Interaction: Natural log of access point volume vs access distance intervals				
50 feet	0.729	0.039	18.597	<0.001
100 feet	0.569	0.131	4.340	<0.001
200 feet	0.358	0.029	12.306	<0.001
300 feet	0.171	0.030	5.629	<0.001
400 feet	0.155	0.021	7.435	<0.001
500 feet	0.116	0.022	5.199	<0.001
600 feet	0.097	0.023	4.299	<0.001
600-1000 feet	0.070	0.017	4.178	<0.001
Overdispersion parameter	0.601	0.057		

Ultimately, the results presented in Table 43 provide guidance as to appropriate distance thresholds for access point location, due to distance threshold indicator variables. Speaking generally, when access points are located further away from the ramp bifurcation point, crash rate is reduced. This affect is first noticeable at 1000 feet from the ramp intersection. The effect is relatively consistent between 1000 feet and 600 feet, with a noticeable increasing trend as the distance becomes less than 600 feet.

4.6 Conclusions

The results of the analysis of the distance between ramp intersections and the nearest access point provide valuable guidance for road agencies. Perhaps the most valuable contribution of this project to the extant literature is the incorporation of access point volume into the functional form of the crash prediction model. Several previous studies have used loosely defined, arbitrary terms such as “major intersection” and “minor intersection” to define the access points in lieu of volume metrics (Rakha et al. 2008, Rakha et al. 2010), however, the framework utilized in this approach provides road agencies with a more flexible, robust basis for decisions as it relates to granting roadway access to the crossroads adjacent to ramp intersections. Model results were intuitive in that four-leg access points and left-turn access points were more dangerous in comparison to the base case of access points on the right side only (as a vehicle travels away from the freeway). Additionally, signalized ramp terminals and signalized access points were associated with higher crash rates on the corridors. Traffic volume on the crossroad was shown to play a significant role in the crash rate on a per mile basis, while the relatively small coefficient for traffic volume from the access point is indicative access point presence, regardless of the volume of vehicles, is associated with reduced safety performance.

These findings provide road agencies and transportation safety researchers with useful information pertaining to the safety of roadway corridors adjacent to freeway ramp intersections. This paper demonstrates that will almost always be adversely impacted by the introduction of access points along a corridor, these effects can be mitigated by locating access points as far from the ramp bifurcation point as possible. Crash rate is significantly impacted when the distance between the ramp terminal and access point is less than 1000 feet, with the most pronounced effects being observed when the distance is 600 feet or less.

CHAPTER 5. SUMMARY

This document investigated traffic safety in three specific areas: vehicular crashes at intersections, pedestrian and cyclist crashes at intersections, and crashes on corridors adjacent to ramp terminal intersections. The results documented in the preceding chapters provide useful insight into approaches to effectively evaluate safety at these facility types.

5.1 Safety Performance Functions for Vehicular Crashes

SPFs for four intersection types (three-leg signalized, three-leg stop-controlled, four-leg signalized, and four-leg stop-controlled) were developed at a variety of complexity levels. In order to accomplish this task, a dataset containing a wide range of geometric and operational characteristics was developed. Three types of "fully specified" negative binomial SPFs: five-year aggregated, five-year aggregated with regional indicators, and a five-year disaggregated random intercept, were estimated and compared to volume-only and volume with regional indicator models. Ultimately this study provides researchers with information that is useful in ascertaining some of the pros and cons of various modeling frameworks and approaches. The series of models estimated in this study provide documentation that the inclusion of regional indicators provides insight on location-specific phenomena associated with crash frequency without overriding the potential to make inferences about various geometric characteristics for this data set. The regional indicators utilized in this study represent the Michigan Department of Transportation's operational/maintenance regions. These seven regions are relatively diverse in terms of traffic volume, which is largely reflective of the underlying population density of each region. The Metro region (which includes Detroit and its suburbs) was used as a basis to which the other regions were compared. Typically, the regional indicators for the non-Metro regions had positive coefficients associated with them. This is largely reflective of two things: the non-linear relationship between traffic volume and crashes at intersections (where the rate at which traffic crashes increase decreases as volume increases) and generally higher volumes in the Metro region. Simultaneously considering the volume-crash relationship and the higher traffic volumes of Metro region explains why for a specific volume, other regions would expect more crashes. This finding is supported by the fact that proximity to DNR non-motorized paths (a variable that is indicative of lower traffic volumes) was associated with increased crash frequency. In addition to the effects of volume captured by the regional indicators, several of the regions (notably the Superior, North, and Southwest) experience relatively severe winters in terms of snow fall which likely also plays a role in the effects of the regional indicators.

In terms of the wide array of geometric characteristics considered in each of the models, the fully specified models possess a great deal of similarity between the coefficients estimated in the five-year aggregated, five-year aggregated with regional indicators, and 5-year disaggregated random intercept model framework.

Generally speaking, the fully-specified models with regional indicators outperform the fully specified models without regional indicators in terms of the model goodness of fit metrics Mean Absolute Deviance (MAD) and Mean Square Predictive Error (MSPE). When the site-specific intercept terms were excluded from the crash prediction estimates, the random-effect models performed similarly to the aggregated five-year models, however, the models dramatically outperformed

Finally, the efforts put provides support for selecting a modeling framework for SPF development if the end goal is to utilize empirical-Bayes methodology to account for site specific crash history. Models were estimated for four facility types, each with its own unique data characteristics, most notably sample mean. This research illustrates the relationship between sample mean, the magnitude of the variance of random (effects) intercept negative binomial models, and the associated overdispersion parameter. As sample mean decreases, the amount of variation of the data explained random intercept, quantified in terms of the variance, increases. When the variance associated with the random intercept term increases, which results in a decrease in the overdispersion parameter. Consequently, if a researcher wishes to utilize the empirical-Bayes approach to incorporate crash history into crash prediction, data sets with low sample means will put higher weight on the model results than a comparable aggregated crash model.

Future work on this topic will include exploration of other modeling frameworks, such as the generalized Poisson, which is equipped to handle under- and overdispersion. An additional topic that warrants investigation is a meta-analysis, particularly focusing on studies that initially calibrated the HSM and then found it appropriate to estimate jurisdiction specific models. Such an analysis may provide insight as to specific characteristics that may be associated with discrepancies between specific locations and the data used in the HSM. Given that investigating the predictive ability of the SPFs was developed using the various frameworks was a focus of this study, splitting the data into training and testing subsets to investigate the out of sample predictive ability of the model types used is also an area of interest.

5.2 Safety Performance Functions for Pedestrian and Cyclist Crashes

Pedestrian and cyclist exposure measures are not commonly maintained by road agencies, and as such, it is difficult for researchers to effectively model crashes for these types of road users on a large scale. Subsequently, studies sometimes forgo pedestrian and/or cyclist exposure measures (Lee and Abdel-Aty, 2005) in crash frequency estimation or approach the problem in terms of a risk factor-based analysis (Bergh and Ray 2014, Bergh et al 2015). The HSM provides a rough framework for the estimation of pedestrian activity at an intersection (AASHTO 2010). Various studies in pedestrian and cyclist safety have utilized study-specific counts (Moreno, and Morency; 2013; Strauss, Miranda-Moreno, and Morency; 2014; Strauss, Miranda-Moreno, and Morency, 2015), and estimated data (Grembek et al 2014), and proposed exposure measures (Lee and Abdel-Aty, 2005), however, the methodologies used to capture pedestrian activity in these studies is not necessarily simple for other researchers or safety professionals to replicate. The methodology outlined in Chapter 3 serves to form a template with which road agencies can feasibly begin to evaluate intersections for pedestrian and cyclist safety through incorporation of the American Community Survey. Separate SPFs were estimated for three crash categories: pedestrian crashes, cyclist crashes, and non-motorized crashes. Models for each crash type were estimated for three-leg signalized intersections, four-leg signalized intersections, and combined signalized intersections.

In addition to demonstrating the usefulness of a publicly available data source to improve the crash prediction efforts for pedestrians and cyclists, this section serves to document how pedestrian and cyclist crashes are affected by various infrastructure components, at least at urban and suburban intersections in Michigan. Similarities between the factors affecting pedestrian and cyclist crashes were observed, beginning with the fact that essentially no pedestrian or bicycle crashes occur at intersections without sidewalks. In addition to the volume components that are traditionally expected to have a large influence on the frequency of crashes, this research demonstrates that in general, median presence, one-way roads, increasing skew angle, and increasing census tract age are associated with decreased crash risk while parking and driveways are associated with lower crash risk. Additionally, certain MDOT regions performed differently from the rest of the state in terms of pedestrian and cyclist safety, which is likely reflective of the actual number of pedestrians and cyclists in that area.

This study presents several approaches to modeling traffic crashes involving pedestrians and cyclists which were compared using MAD and MSPE. Two modeling approaches tended to outperform the others: estimating separate models for each crash type (pedestrian and cyclist) and intersection type (three-leg signalized, four-leg signalized), and estimating separate models for each crash type but simultaneously considering each intersection type with the use of an indicator variable. The worst predictive performance of the models estimated for this study considered pedestrian and cyclist crashes combined into non-motorized crashes and then used the proportion of pedestrian commuter density and cyclist commuter density relative to the combined non-motorized road user density to determine the number of predicted crashes for each separate crash type.

This study is limited in that no actual pedestrian or cyclist data was available to serve as a means to verify the effectiveness of the ACS data as exposure measure. Future work in this area could attempt to estimate similar models using observed pedestrian and cyclist volumes, estimated volumes using land use data or other sources, and the ACS data to contrast the differences in results and provide insight into the explanatory and predictive capabilities of each model. Additionally, pedestrian and cyclist crashes could be aggregated at the census tract level, rather than the intersection level to provide another perspective as to the factors affecting these types of crashes.

5.3 Access Point Proximity to Crossroad Ramp Terminals

Corridors adjacent to ramp terminal intersections are a focal point in the access management policies of many road agencies. While the extant literature clearly demonstrates that higher access point density is associated with increased crash risk, the issue of locating an access point near to a ramp terminal intersection is not often approached as a pure access point density issue. This paper substantially adds to the extant literature in this area. Several previous studies have used loosely defined, arbitrary terms such as "major intersection" and "minor intersection" to define the access points in lieu of volume metrics (Rakha et al. 2008, Rakha et al. 2010), however, this is one of the first studies to incorporate access point volume into the analysis framework. The utilization of access point volume estimated based on existing roadway information where available, in conjunction with driveway volume from the ITE Trip Generation Manual in the model instead of access point classifications, which can be arbitrary, provides road agencies with a flexible structure on which they can base their specific access management policy.

The analysis approach utilized in this study was two-pronged. First a thorough visual inspection of the data was undertaken to gain high-level insight into the underlying relationships between traffic volume, distance between the ramp intersection and first access point, and the crash rate (in vehicles per mile). The visual inspection of the data revealed an inverse relationship between crashes per mile and distance between the ramp intersection and access point. When separate traffic volume intervals were examined, an increasing relationship between traffic volume and crash rate became visible, as one would expect.

Following the visual analysis, a detailed statistical analysis was conducted utilizing random effects negative binomial regression models. Model results were intuitive in that four-leg access points and left-turn access points were more dangerous in comparison to the base case of access points on the right side only (as a vehicle travels away from the freeway). Additionally, signalized ramp terminals and signalized access points were associated with higher crash rates on the corridors. In order to facilitate interpretation of the model results, a series of graphical representations of the model were created. These figures demonstrate the relationship that exists between crash rate, traffic volume, and the distance between the ramp terminal intersection and the nearest access point.

This study was somewhat limited in terms of geometric characteristics present on the crossroad, due in part to the rural nature of many interchanges in the state of Iowa. Future work could aim to incorporate data from other, more densely populated areas which may provide better insight as to specific geometric characteristics effecting safety near ramp terminal intersections. Additional information could also be collected for comparable corridors that are not ramp-adjacent. This would allow for an analysis to provide insight as to the safety of these corridors relative to the broader population of roadways in Iowa which could potentially result in the determination of a “base” level of safety to which the ramp-adjacent corridors could be compared. Finally, given that one of the most commonly identified reasons for locating access in close proximity to ramp terminal intersections is loss of potential revenue, an ideal component that could be added to this research in the future would involve approaching the problem from economic standpoint utilizing monetary values associated with crash severity. This approach will allow road agencies and developers to communicate in terms of the same units.

5.4 Conclusion

The research contained in this dissertation represent significant contributions to the body of research literature in traffic safety. The findings help to shed light on several under-researched areas of transportation safety, specifically intersection safety. In addition to the filling gaps in the extant literature, these findings also represent useful approaches to solving problems that researchers and road agencies alike must address. This research accomplishes several tasks, most notably: examining the consequences of selecting various model frameworks for the estimation of crash prediction models, utilizing existing data sources to improve the ability to model pedestrian and cyclist crashes, and providing support to road agencies in addressing access management related to the proximity of ramp terminal intersections and access points. Collectively, this document adds to the extant body of research in these areas in ways that provide utility to road agencies and researchers alike. Additionally, this document provides insight into future work that will continue to add value to transportation safety literature

REFERENCES

Alluri P (2008). *Assessment of Potential Site Selection Methods for Use in Prioritizing Safety Improvements on Georgia Roadways*. MS thesis. Clemson University, Clemson, SC.

Alluri P, Ogle J (2012). Effects of State-Specific SPF's, AADT Estimations, and Overdispersion Parameters on Crash Predictions Using SafetyAnalyst. *Presented at 91st Annual Meeting of the Transportation Research Board, Washington, DC.*

Alluri P, Ogle J (2012). Road Safety Analysis in the United States: States' Current Practices and Their Future Direction. *Transportation Research Record: Journal of the Transportation Research Board, No. 2318, Transportation Research Board of the National Academies*, Washington, DC, 7–15.

Alluri P, Saha D, Liu K, Gan A (2014). *Improved Processes for Meeting the Data Requirements for Implementing the Highway Safety Manual (HSM) and SafetyAnalyst in Florida*. Publication BDK80-977-37, Florida Department of Transportation, Tallahassee, FL.

American Association of State Highway and Transportation Officials (AASHTO) (1991). *A Policy on Geometric Design of Highways and Streets,* Washington DC.

American Association of State Highway and Transportation Officials (AASHTO) (2004). *A Policy on Geometric Design of Highways and Streets,* Washington DC.

American Association of State Highway and Transportation Officials (AASHTO) (2010). *Highway Safety Manual, First Edition*, Washington, DC.

American Association of State Highway and Transportation Officials (AASHTO). *An Introduction to the Highway Safety Manual*. http://www.highwaysafetymanual.org/Documents/HSMP-1.pdf, Accessed on October 29, 2012.

American Association of State Highway and Transportation Officials (AASHTO). *Highway Safety Manual Frequently Asked Questions.* http://www.highwaysafetymanual.org/Pages/support_answers.aspx#43. Accessed April, 2015.

American Association of State Highway and Transportation Officials (AASHTO). *SafetyAnalyst Overview*. http://www.safetyanalyst.org/index.htm. Accessed April 2015.

Barbosa H, Cunto F, Bezerra B, Nodari C, Jacques MA (2014). Safety Performance Models for Urban Intersections in Brazil. *Accident Analysis and Prevention*, Vol. 70, 258-266.

Bennet SA, Yiannakoulias N (2015). Motor-vehicle collisions involving child pedestrians at intersection and mid-block locations. *Accident Analysis and Prevention*, Vol 78, Elsevier, 94-103.

Bergh C, Ray B (2014). *ODOT Pedestrian and Bicycle Safety Implementation Plan.* Project No. 13452.0, Oregon Department of Transportation.

Bergh C, Griffin A, Braghton M, Boudart J (2015). Implementing a Risk-Based Safety Prioritization Method for Pedestrian and Bicycle Crashes in Oregon. *Transportation Research Board 94th Annual Meeting Compendium of Papers,* Washington, DC.

Bonneson JA, Geedipally S, Pratt MP, Lord D (2012). *National Cooperative Highway Research Program: Safety Prediction Methodology and Analysis Tool for Freeways and Interchanges*, Publication NCHRP 17-45, NCHRP, Transportation Research Board, Washington, DC.

Bornheimer C, Schrock S, Wang M, Lubliner H (2012). Developing a Regional Safety Performance Function for Rural Two-Lane Highways. *Presented at 91st Annual Meeting of the Transportation Research Board*, Washington, DC.

Brimley BK, Saito M, Schultz GG (2012). Calibration of the Highway Safety Manual Safety Performance Function: Development of New Models for Rural Two-Lane Two-Way Highways. *Transportation Research Board: Journal of the Transportation Research Board*, No. 2279, Transportation Research Board of the National Academies, Washington, DC.

Butorac MA, Wen JC (2004). *National Cooperative Highway Research Program Synthesis 332 Access Management on Crossroads in the Vicinity of Interchanges: A Synthesis of Highway Practice*. Transportation Research Board, Washington DC.

Cafiso S, Di Silvestro G, Di Guardo G (2012). Application of Highway Safety Manual to Italian Divided Multilane Highways. *Procedia-Social and Behavioral Sciences*, Vol. 53, 911-920.

Chen Y, Persaud B, Sacchi E (2012). Improving Transferability of Safety Performance Functions by Bayesian Model Averaging. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2280, Transportation Research Board of the National Academies, Washington, DC, 162–172.

Cunto F, Sobreira L, Ferreira S (2015). Assessing the Transferability of the Highway Safety Manual Predictive Method for Urban Roads in Fortaleza City, Brazil. *Journal of Transportation Engineering*, Vol. 141(1), 8p.

Dai D (2012). Identifying clusters and risk factors of injuries in pedestrian-vehicle crashes in a GIS environment. *Journal of Transport Geography,* vol. 24, Elsevier, 206-214.

Dixon K, Monsere C, Xie F, Gladhill K (2012). *Calibrating the Future Highway Safety Manual Predicitive Methods for Oregon State Highways*. Publication FHWA-OR-RD-12-07. Oregon Department of Transportation, Salem, OR, Federal Highway Administration, Washington, DC.

Donnell E, Gayah V, Jovanis P (2014). *Safety Performance Functions*. Publication FHWA-PA-2014-007-PSU WO 1. Pennsylvania Department of Transportation, Bureau of Planning and Research, Harrisburg, PA.

Ewing R, Dumbaugh E (2009). The Built Environment and Traffic Safety. *Journal of Planning Literature*, Vol. 23, No. 4, 247-267.

Fixing America's Surface Transportation (FAST) Act of 2015 Pub. L. No. 114-94.

Florida Department of Transportation (2014). *Median Handbook. Florida Department of Transportation, Tallahassee, FL.*

Garber NJ, Haas PR, Gosse C (2010). *Development of Safety Performance Functions for Two-Lane Roads Maintained by the Virginia Department of Transportation*. Publication FHWA/VTRC 10-R25. FHWA, Virginia Department of Transportation, Richmond, VA.

Garber NJ, Rivera G (2010). *Safety Performance Functions for Intersections on Highways Maintained by the Virginia Department of Transportation*. Publication FHWA/VTRC 11-CR1. FHWA, Virginia Department of Transportation, Richmond, VA.

Gitelman V, Balasha D, Carmel R, Hendel L, Pesahov F (2012). Characterization of pedestrian accidents and an examination of infrastructure measures to improve pedestrian safety in Israel. *Accident Analysis and Prevention*, vol 44, Elsevier, 63-73.

Giuffre O, Grana A, Giuffre T, Marino R, Marino S (2014). Estimating the safety performance function for urban unsignalized four-legged one-way intersections in Palermo, Italy. *Archives of Civil Engineering*, Vol. 60(1), 41-54.

Gluck J, Levinson HS, Stover V (1999). *National Cooperative Highway Research Program Synthesis 404: State of the Practice in Highway Access Management*. Transportation Research Board, Washington DC.

Gluck J, Lorenz MR (2010). *National Cooperative Highway Research Program Report 420: Impacts of Access Management Techniques*. Transportation Research Board, Washington DC.

Grembek O, Bosman C, Bigham JM, Fine S, Griswold JB, Medury A, Sanders RL, Schneider RJ, Yavari A, Zhang Y, Ragland DR. *Pedestrian Safety Improvement Program*. Report No. CA14-2209, California Department of Transportation, Sacramento CA.

Harwood DW, Bauer KM, Richard KR, Gilmore DK, Graham JL, Potts IB, Torbic DJ, Hauer E (2007). *National Cooperative Highway Research Program Document 129, Phase I and II: Methodology to Predict the Safety Performance of Urban and Suburban Arterials*, Publication NCHRP 17-26, NCHRP, Transportation Research Board, Washington, DC.

Harwood DW, Council FM, Hauer E, Hughes WE, Vogt A (2000). *Prediction of the Expected Safety Performance of Rural Two-Lane Highways*. Publication FHWA-RD 99-207. FHWA, U.S. Department of Transportation, McLean, VA.

Harwood DW, Torbic DJ, Gilmore DK, Bokenkroger CD, Dunn JM, Zegeer CV, Srinivasan R, Carter D, Raborn C (2008). *National Cooperative Highway Research Program Document 129, Phase III: Methodology to Predict the Safety Performance of Urban and Suburban Arterials: Pedestrian Safety Prediction Methodology*, Publication NCHRP 17-26, NCHRP, Transportation Research Board, Washington DC.

Hauer E (1997). *Observational Before–After Studies in Road Safety*. Pergamon Press, Elmsford, NY.

Hauer E, Kononov J, Allery B, Griffith MS (2002). Screening the Road Network for Sites with Promise. *Transportation Research Record: Journal of the Transportation Research Board*, No. 1784, Transportation Research Board of the National Academies, Washington, DC,27-32. 5.

Institute of Transportation Engineers (ITE) (2008). *Trip Generation Manual, 8th Edition*. Washington DC.

Kim D (2009). Improvement of Pedestrian and Bicycle Safety: Addressing Environmental Contributing Factors. *Korean Local Administration Review,* Vol. 6, No. 1,

Kweon, Y, Lim I, Turpin, TL, Read SW (2014). Development and Application of Guidance to Determine the Best Way to Customize the Highway Safety Manual for Virginia. In Transportation Research Record: Journal of Transportation Research Board, No. 2435, Transportation Research Board of the National Academies, Washington, DC,. 27-36.

Kweon, Y, Lim I (2014). Development of Safety Performance Functions for Multilane Highway and Freeway Segments Maintained by the Virginia Department of Transportation. Report No. FHWA/VCTIR 14-R14. Virginia Department of Transportation and the Federal Highway Administration, Richmond, VA.

Lee C, Abdel-Aty M (2005). Comprehensive analysis of vehicle-pedestrian crashes at intersections in Florida. *Accident Analysis and Prevention,* vol 37, Elsevier, 775-786.

Lord D (2006). Modeling Motor Vehicle Crashes Using Poisson-gamma Models: Examining the Effects of Low Sample Mean Values and Small Sample Size on the Estimation of the Fixed Dispersion Parameter. *Accident Analysis and Prevention*, Vol. 38(4), 751-766.

Lord D, Mannering F (2010). The Statistical Analysis of Crash-Frequency: A Review and Assessment of Methodologies. *Transportation Research, Part A: Policy and Practice,* Vol. 44, Iss. 5, 291-305.

Lord D (2016). *National Cooperative Highway Research Program: Safety Prediction Models for Six-Lane and One-Way Urban and Suburban Arterials*. Publication NCHRP 17-58, NCHRP, Transportation Research Board, Washington, DC (Forthcoming).

Lord D, Geedipally SR, Persaud BN, Washington SP, van Schalkwyk I, Ivan JN, Lyon C, Jonsson T (2008). *National Cooperative Highway Research Program Document 126: Methodology for Estimating the Safety Performance of Multilane Rural Highways*. Publication NCHRP 17-29, NCHRP, Transportation Research Board, Washington, DC.

Lord D, Park PY (2008). Investigating the Effects of the Fixed and Varying Dispersion Parameters of Poisson-gamma Models on Empirical Bayes Estimates. *Accident Analysis and Prevention*, Vol. 40, 1441-1457.

Lu J, Gan A, Haleem K, Alluri P, Liu K (2012). Comparing Locally-Calibrated and SafetyAnalyst-Default Safety Performance Functions for Florida's Urban Freeways. *Presented at 91st Annual Meeting of the Transportation Research Board*, Washington, DC.

Lubliner H, Bornheimer C, Schrock S, Wang MH, Fitzsimmons E (2014). *Evaluation of Interactive Highway Safety Design Model Crash Prediction Tools for Two-Lane Rural Roads on Kansas Transportation Projects*. Publication KU-10-1R, Kansas Department of Transportation, Topeka, KS.

Lubliner H, Schrock SD (2012). Calibration of the Highway Safety Manual Prediction Method for Rural Kansas Highways. *Presented at 91st Annual Meeting of the Transportation Research Board*, Washington, DC.

Lyon C, Gotts B, Wong WKF, Persaud B (2007). Comparison of Alternative Methods for Identifying Sites with High Proportion of Specific Accident Types. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2019, Transportation Research Board of the National Academies, Washington, DC, 2007, 212–218.

Martinelli F, La Torre F, Vadi P (2009). Calibration of the Highway Safety Manual's Accident Prediction Model for Italian Secondary Road Network. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2103, Transportation Research Board of the National Academies, Washington, DC, 1–9.

Monsere CM, Johnson T, Dixon K, Zheng J, van Schalkwyk I (2011). *Assessment of Statewide Intersection Safety Performance*. Publication FHWA-OR-RD-18. Oregon Department of Transportation Research Section, Salem OR and FHWA, Washington, DC.

Moving Ahead for Progress in the 21st Century Act of 2012, Public L. No. 112-141.

National Highway Traffic Safety Administration's Center for Statistics and Analysis (2009A). *Traffic Safety Facts: 2009 Data-Pedestrians.* NHTSA, Washington, DC.

National Highway Traffic Safety Administration's Center for Statistics and Analysis (2009A). *Traffic Safety Facts: 2009 Data-Bicyclists and Other Cyclists.* NHTSA, Washington, DC.

National Highway Traffic Safety Administration (NHTSA). *About MMUCC*. http://www.mmucc.us/about-mmucc. Accessed April, 2015.

Oh J, Lyon C, Washington S, Persaud B, Bared J (2003). Validation of FHWA crash models for rural intersections: lessons learned. *Transportation Research Record: The Journal of the Transportation Research Board of the National Academies,* 1840, Washington DC, 41-49.

Persaud B, Lyon C (2009). *Safety Performance Functions for Intersections*. Publication CDOT-2009-10. Colorado Department of Transportation, Denver, CO.

Persaud B, Nguyen T (1998). *Disaggregate Safety Performance Models for Signalized Intersections on Ontario Provincial Roads*. Transportation Research Record: Journal of the Transportation Research Board, No. 1635, Transportation Research Board of the National Academies, Washington, DC, 113-120.

Persaud BN (2001). *NCHRP Synthesis of Highway Practice 295: Statistical Methods in Highway Safety Analysis*. TRB, National Research Council, Washington, DC.

Prato CG, Gitelman V, Bekhor S (2012). Mapping patterns of pedestrian fatal accidents in Israel. *Accident Analysis and Prevention, Vol. 44, Elsevier, 56-62.*

Pulugurtha SS, Sambhara VR (2011). Pedestrian crash estimation models for signalized intersections. *Accident Analysis and Prevention,* Vol 43, Elsevier, 439-446.

Rakha H, Flintsch AM, Arafeh M, Abdel-Salam G, Dua D, Abbas M (2008). *Access Control Design on Highway Interchanges.* Report No. VTRC 08-CR7. Virginia Department of Transportation, Richmond VA.

Rakha H, Arafeh M, Abdel-Salam G, Guo F, Flintsch AM, (2010). *Linear Regression Crash Prediction Models: Issues and Proposed Solutions.* Report No. VT-2008-02. Virginia Department of Transportation, Richmond VA, U.S. Department of Transportation, Washington DC.

Sacchi E, Persaud B, Bassani M (2012). Assessing International Transferability of Highway Safety Manual Crash Prediction Algorithm and Its Components. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2279, Transportation Research Board of the National Academies, Washington, DC, 90–98.

Savolainen PT, Gates TJ, Lord D, Geedipally S, Rista E, Barrette TP, Russo BJ, Hamzeie R (2015). *Michigan Urban Trunkline Intersections Safety Performance Functions (SPFs) Development and Support*. Report No. RC-1628. Michigan Department of Transportation, Lansing, MI.

Shin H, Lee Y-J, Dadvar S (2014). *The Development of Local Calibration Factors for Implementing the Highway Safety Manual in Maryland*. Publication SP209B4J, Office of Traffic and Safety (OOTS), Maryland State Highway Administration, Baltimore, MD.

Srinivasan R, Carter, D, Bauer K (2013). *Safety Performance Function Decision Guide: SPF Calibration versus SPF Development*. Report No. FHWA-SA-14-004, Federal Highway Administration, Washington, DC.

Srinivasan R, Carter D (2011). *Development of Safety Performance Functions for North Carolina*. Publication FHWA/NC/2010-09. Federal Highway Administration Office of Safety, Washington, DC.

Srinivasan R, Council FM, Harkey DL (2008). *Calibration Factors for HSM Part C Predictive Models*. Unpublished memorandum prepared as part of the Federal Highway Administration Highway Safety Information System project, Highway Safety Research Center, University of North Carolina, Chapel Hill, NC.

Srinivasan S, Haas P, Dhakar NS, Hormel R, Torbic D, Harwood D (2011). *Development and Calibration of Highway Safety Manual Equations for Florida Conditions*. Publication TRC-FDOT-82013-2011. Florida Department of Transportation, Tallahassee, FL.

Strauss J, Miranda-Moreno LF, Morency P (2013). Cyclist activity and injury risk at signalized intersections: A Bayesian modelling approach. *Accident Analysis and Prevention*, vol 59, Elsevier, 9-17.

Strauss J, Miranda-Moreno LF, Morency P (2015). Mapping cyclist activity and injury risk in a network combining smartphone GPS data and bicycle counts. *Accident Analysis and Prevention*, Vol 83, Elsevier, 132-142.

Sun C, Brown H, Praveen E, Claros B, Nam K (2014*). Calibration of the Highway Safety Manual for Missouri*. Publication TR201302. Missouri Department of Transportation, Jefferson City, MO.

Sun R, Zhuang X, Wu C, Zhao G, Zhang K (2015). The estimation of vehicle speed and stopping distance by pedestrians crossing streets in a naturalistic traffic environment. *Transportation Research Part F*, vol 30, Elsevier, 97-106.

Tegge RA, Jo J, Ouyang Y (2010). *Development and Application of Safety Performance Functions for Illinois*. Publication FHWA-ICT-10-066. Illinois Department of Transportation, Springfield, IL.

Tolford T, Renne J, Fields B (2014). Development of Low-Cost Methodology for Evaluating Pedestrian Safety in Support of Complete Streets Policy Implementation. *Transportation Research Record: Journal of the Transportation Research Board,* No. 2464, Transportation Research Board of the National Academies, Washington DC, 29-37.

Transportation Research Board (TRB) (2010). Highway Capacity Manual. National Research Council, Washington, DC.

Troyer D, Bradbury K, Juliano C (2015). Strength of the Variable: Calculating and Evaluating Safety Performance Function Calibration Factors for the State of Ohio. *Presented at 94th Annual Meeting of the Transportation Research Board*, Washington, DC.

University of North Carolina Safety Research Center. Crash Modification Factor Clearinghouse. http://www.cmfclearinghouse.org/. Federal Highway Administration, Accessed on May 21, 2015.

Vandenbulcke G, Thomas I, Int Panis L, (2014). Predicting cycling accident risk in Brussels: A spatial case-control approach. *Accident Analysis and Prevention,* Vol 62, Elsevier, 341-257.

Vogt A, Bared JG (1998). *Accident Models for Two-Lane Rural Roads: Segments and Intersections.* Publication FHWA-RD-98-133. FHWA, U.S. Department of Transportation, McLean, VA.

Vogt, A (1999). *Crash Models for Rural Intersections: 4-Lane by 2-Lane Stop-Controlled and 2-Lane by 2-Lane Signalized*. Publication FHWA-RD-99-128. FHWA, U.S. Department of Transportation, McLean, VA.

Wang H., Chitturi M, Noyce DA (2011). Developing Safety Performance Functions for Diamond Interchange Ramp Terminals. *Presented at 90th Annual Meeting of the Transportation Research Board,* Washington, DC.

Wang Y, Kockelman KM (2013). A Poisson-lognormal conditional-autoregressive model for multivariate spatial analysis of pedestrian crash counts across neighborhoods. *Accident Analysis and Prevention,* vol 60, Elsevier, 71-84.

Wang Y, Sharda S, Wang H (2016). A Systemic Safety Analysis of Pedestrian Crashes: Lessons Learned. *Transportation Research Board 95th Annual Meeting Compendium of Papers*, Washington DC.

Washburn SS, Kondyli A (2006). *Development of Guidelines for Driveway Location and Median Configuration in the Vicinity of Interchanges.* Report No. TRC-FDOT-036-2006, Florida Department of Transportation, Tallahassee, FL.

Williams KM (2003). Access Management Manual: TRB Committee Documents State of the Art. TR News, 228, Washington, DC, 39-40.

Williams KM, Zhou H, Hagen L, Farah W (2004). Benefit-and-Cost Analysis of Strategic Acquisition of Limited Access Right-of-Way near Interchanges. *6th National Access Management Conference*, Kansas City, MO.

World Health Organization (2013). *Global Status Report on Road Safety 2013: Supporting a Decade of Action.* World Health Organization, Geneva, Switzerland.

Young J, Park P (2012). Comparing the Highway Safety Manual's Safety Performance Functions with Jurisdiction-Specific Functions for Intersections in Regina. *Implementation of Safety Prediction Methods in the Processes and Procedures of a Transportation Agency Session of the 2012 Annual Conference of the Transportation Association of Canada Fredericton*, New Brunswick, Canada.

Zhao J, Liu Y (2016). Integrated signal optimization and non-traditional lane assignment for urban freeway off-ramp congestion mitigation. *Transportation Research Part C*, 73, Elsevier, 219-238.

APPENDIX: DISAGGREGATE-LEVEL ASSESSMENT OF CHANGES TO MICHIGAN'S MOTORCYCLE HELMET USE LAW

Effects on Motorcyclist Injury Outcomes

Timothy P. Barrette, Trevor J. Kirsch, Peter T. Savolainen, Brendan J. Russo, and Timothy J. Gates

Published in: Transportation Research Record: Journal of the Transportation Research Board, No. 2468, Transportation Research Board of the National Academies, Washington, D.C., 2014, pp. 131-137.

Abstract

On April 13, 2012, Michigan instituted reforms to its long-standing universal helmet law (UHL) and thus became the 28th state with a partial helmet law, which requires helmet use by only a subset of the riding population. Given continuing increases in motorcycle fatalities, helmet use remains a divisive policy issue facing state governments. The existing research literature includes several before-and-after studies that evaluate the effects of changes in motorcycle helmet laws on metrics such as helmet use and statewide motorcycle fatalities. However, a comprehensive assessment of the effects of helmet use laws on the full range of crash injury outcomes is lacking. Important evidence was added to inform the continuing debate about the efficacy of UHLs. A detailed, disaggregate-level study was conducted to assess the degree of injury severity sustained by motorcyclists involved in crashes before and after Michigan's transition from a UHL to a partial helmet law. By control- ling for various rider, roadway, traffic, and weather characteristics, results of the study demonstrate that helmets reduced the probability of fatalities by more than 50%. Injuries tended to be less severe in crashes that involved deer, occurred at lower speeds, or occurred in inclement weather but more severe in high-speed collisions or when drugs and alcohol were involved. The riders who tended to be more susceptible to severe injury were female, younger (age 21 to 30 years), or older (age 51 to 70 years).

Introduction

On April 13, 2012, the Michigan legislature amended Section 658 of the Michigan Vehicle Code (PA 300 of 1949) and weakened a universal helmet law (UHL) for motorcycles that had been in place since 1969. The resultant partial helmet law (PHL) allows motor- cycle operators more than 21 years of age to ride without a helmet if they have had a motorcycle endorsement for at least 2 years or have passed a motorcycle safety course and carry $20,000 of insurance per person on the motorcycle. This legislative change made Michigan one of 28 states with a PHL, which requires helmet use by only a subset of the riding population (*1*). As of August 2013, only 19 states have UHLs and three have no helmet laws (*1*). On a map of helmet laws by state, Figure 1 shows that UHLs are predominantly in effect along the East Coast and the West Coast.

Helmet use laws have been an issue of considerable debate among the motorcycle community and the general public, even though helmets have been shown to be 29% effective in preventing motorcycle fatalities and 67% effective in preventing brain injuries resulting from motorcycle crashes (*2*). Research also has shown that riders who do not wear helmets also are more likely to require a skilled nursing facility, and various studies have shown that fatality rates are lower in states with UHLs (*3–5*). The existing research literature includes numerous studies that have examined the effects of motor- cycle helmet use and UHLs. The National Highway Traffic Safety Administration estimates that 17,572 lives were saved by motor-cycle helmets from 1984 to 2005, and an additional 11,568 could have been saved by helmet use (*6*). Research has demonstrated that helmets save lives and that helmeted riders have lower hospitalization rates and are 2.4 times less likely to suffer a head injury (*7, 8*). Furthermore, research has shown that after a repeal of helmet laws, the number of drinking-related and unhelmeted fatal crashes increases significantly (*9*). Extensive research also has been conducted to examine how helmet laws affect helmet use, because rates can be more than 90% in UHL states versus rates as low as 55% in PHL states (*4*).

Collectively, these findings are consistent with those of many prior studies, which have shown conclusive benefits of helmet use and UHLs (*10–22*). In fact, PHLs have been suggested to be essentially equivalent to a full repeal because of the difficulty of enforcing violations, which usually are based on age and experience (*11*).

However, some studies have drawn conflicting conclusions about UHL efficacy. For example, Stolzenberg and D'Alessio found no significant difference in the rate of fatal brain injuries 18 months after Florida repealed its UHL in 2000 (*23*). They also noted that various prior studies had failed to consider temporal trends appropriately. Other research has shown that the impacts of helmet laws may be understated if appropriate controls (e.g., temperature and weather) are not accounted for (*5*).

Opponents of helmet use legislation frequently cite potential economic benefits that increased tourism would generate as a result of fewer riding restrictions. However, a 2012 study estimated that the weakening of Michigan's helmet use law would result in increases of 42% in monetary costs and 54% in non- monetary costs resulting from motorcycle-involved crashes (*24*). On average, initial medical costs were \$5,000 higher for unhelmeted riders than for helmeted riders (*25*); after the weakening of Florida's motorcycle helmet law, the medical costs of motorcyclists being admitted to hospitals with head, brain, or skull injuries more than doubled from \$21 million to \$44 million (*26*). Compounding this rise was a parallel increase in treatment costs for such injuries from \$34,518 to roughly \$40,000.

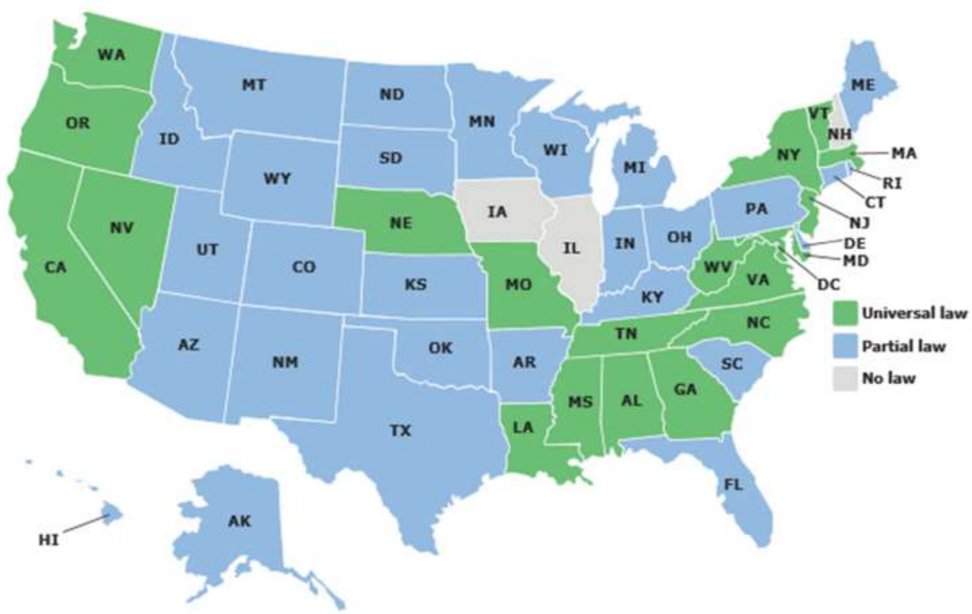


FIGURE 1 Motorcycle helmet laws, by state (AK = Alaska; AL = Alabama; AR = Arkansas; AZ = Arizona; CA = California; CO = Colorado; CT = Connecticut; DC = District of Columbia; DE = Delaware; FL = Florida; GA = Georgia; HI = Hawaii; IA = Iowa; ID = Idaho; IL = Illinois; IN = Indiana; KS = Kansas; KY = Kentucky; LA = Louisiana; MA = Massachusetts; MD = Maryland; ME = Maine; MI = Michigan; MN = Minnesota; MO = Missouri; MS = Mississippi; MT = Montana; NC = North Carolina; ND = North Dakota; NE = Nebraska; NH = New Hampshire; NJ = New Jersey; NM = New Mexico; NV = Nevada; NY = New York; OH = Ohio; OK = Oklahoma; OR = Oregon; PA = Pennsylvania; RI = Rhode Island; SC = South Carolina; SD = South Dakota; TN = Tennessee; TX = Texas; UT = Utah; VA = Virginia; VT = Vermont; WA = Washington State; WV = West Virginia; WI = Wisconsin; WY = Wyoming).

Despite these economic benefits, which are well supported, debate continues about the efficacy of helmet use laws in reducing motor- cycle fatalities (*27–32*). The principal objective of this study was to determine the impacts of the recent weakening of Michigan's UHL on injury outcomes in crashes that involve motorcycles. A detailed, disaggregate-level assessment was developed for motorcyclist crash injury outcomes before and after the weakening of Michigan's helmet use law. A random effects ordered probit model was estimated to ascertain the effects of helmet use while controlling for other important factors. Collectively, the results provide important evidence to guide subsequent policy decisions in Michigan and other states.

To assess the effects of helmet use on the degree of injury sustained as a result of motorcycle crashes in the state of Michigan, an ordered probit model was developed. The ordered probit is an appealing analytical framework in that it accounts for the ordinal nature of injury data, which can be ranked in ascending order of severity from property damage only (no injury) to fatal injury. For the ordered probit model, a latent variable (z) is specified as a linear function for each crash observation (33, 34), such that

$$z = \beta X + \varepsilon \tag{1}$$

where,

β = vector of estimable parameters,
X = vector of variables determining discrete ordering for each crash observation, and
ε = disturbance term

The, observed ordinal-injury data (y) for each observed crash are defines as

$$y = \begin{cases} 1 & if\ z \leq \mu_0 \\ 2 & if\ \mu_0 < z \leq \mu_1 \\ 3 & if\ \mu_1 < z \leq \mu_2 \\ \ldots & \\ i & if\ z > \mu_{i-1} \end{cases} \tag{2}$$

where μ are estimable threshold parameters that define y (which corresponds to integer ordering) and i is the highest integer-ordered response. The μ parameters are estimated jointly with the model parameters β, and μ0 can be set to 0 without loss of generality. The estimation problem then becomes one of determining the probability of i-specific ordered responses for each crash injury (n). If the error term ε is assumed to be normally distributed across observations with a mean of 0 and variance of 1, then an ordered probit model results.

Setting the lower threshold μ_0 equal to 0 results in the outcome probabilities

$$P(y = 1) = \Phi(\mu_i - \beta X) - \Phi(\mu_{i-1} - \beta X) \tag{3}$$

where Φ equals the standard normal cumulative density function; and μ_i and μ_{i-1} represent the upper and lower thresholds, respectively, for injury severity *i.* For the purposes of this study, one potential concern in analyzing injury severity data is that the rider and pillion passenger on the same motorcycle probably share common unobserved effects, for example, being exposed to similar impact forces, wearing similar gear, or sharing other similarities that cannot be captured by the available data from the crash report form. Failure to account for this correlation potentially can result in inefficient or biased parameter estimates. To address this issue, a motorcycle-specific disturbance term (φ_i) is added to account for the random unobserved effects z that are specific to each crash-involved motorcycle c as follows:

$$z_{ic} = \beta X_{ic} + \varepsilon_{ic} + \varphi_i \quad (4)$$

The random effects ordered probit model is estimated by standard maximum likelihood methods.

Data Summary

Summary statistics for all motorcycle crashes occurring in the state of Michigan over two periods (April 13 to December 31, 2011, and April 13 to December 31, 2012) are listed in Table 1. The April 13 date coincides with enactment of the PHL and, as such, enables a direct comparison of changes in motorcycle crashes before and after the helmet law change.

During this period in 2011(when the UHL was in place), 2,979 motorcycles were involved in crashes in the state of Michigan, of which 105 resulted in motorcyclist fatalities. During the same period in 2012 (under the PHL), motorcycle crashes increased by 6.3% to 3,166 and fatalities increased by 11.4% to 117; incapacitating injuries also increased by 11.2% (from 544 to 605), thus providing general data that crash severity increased after the helmet law change.

Over the 2011 period, the helmet use rate among all crash-involved motorcyclists in Michigan was 94.3%. During the same period in 2012, the rate dropped to 72.5%. The latter rate is close to the estimated statewide use rate of 73.0% determined in a recent direct observation survey (35). The variables in Table 1 other than helmet use and injury severity were largely consistent between the two periods.

Results and Discussion

Results of the random effects ordered probit model that was estimated as part of this study are presented in Table 2. They show that various motorcyclist, crash, roadway, and temporal factors influence the severity of injuries sustained as a result of motorcycle crashes. In interpreting model results from Table 2, a positive coefficient implies that as that variable is changed from 0 to 1, the probability of a fatal injury increases while the probability of the motorcyclist sustaining no injury decreases (and vice versa for negative coefficients).

The interpretation of changes on interior injury categories is not intuitive but requires the calculation of elasticities. As presented in Table 3, elasticities indicate the percentage of change in the probability of each injury outcome as each variable is increased from 0 to 1. In practical terms, these values represent the percentage change in the probability of a specific severity level due to the effects of a specific indicator variable. For example, the results indicate that the probability that a crash results in property damage only (O) increases by 27.2% when a rider is wearing a helmet (versus not wearing a helmet); similarly, the probability of a possible (C) injury increases by 10.7% when a helmet is used. Conversely, the likelihood of nonincapacitating (B), incapacitating (A), and fatal (K) injuries decrease by 5.4%, 26.1%, and 51.6%, respectively. Collectively, these findings provide additional evidence to demonstrate that helmet use leads to consistent, pronounced reductions in injury severity.

Including helmet use, a total of 22 explanatory variables significantly affect the injury severity outcomes of motorcyclists involved in a crash. Age has a significant influence on the severity of injuries sustained by motorcyclists. The model shows that motorcycle riders from 21 to 30 years of age and from 51 to 70 years of age had similar injury outcome characteristics and were more likely to experience severe crash-related injuries. The younger motorcyclists may be inherently riskier drivers and engage in reckless behavior while riding, and the finding that they are at greater risk of injury is consistent with past studies (20). Older drivers may be at greater risk of injury or fatality due to the effects of aging, such as reduced reaction time or frailty; this finding also is consistent with past findings (36). Gender also had an effect on injury outcomes because females were at greater risk of severe or fatal injury.

Alcohol use and drug use significantly increase the likelihood of a motorcyclist sustaining a fatal injury (by 112.9% and 237%, respectively). These results are not surprising, because alcohol and drug use can affect cognitive abilities in many ways, potentially leading to slower reaction times, poor judgment, and a false sense of confidence. These findings are consistent with previous studies and strengthen the argument for continuing education and enforcement campaigns aimed at reducing impaired riding (20).

Riders involved in collisions with a deer were less likely to experience fatal injuries and more likely to experience no injury, probably because the mass of a deer is smaller than that of a motor vehicle. Motorcyclists involved in collisions with large trucks were 208.9% more likely to experience a fatal injury. This finding is not surprising because of the sheer size and mass differential between a motorcycle and a large truck.

Crash type also significantly affected injury severity levels of crash-involved motorcyclists. Riders involved in single-vehicle, rear-end, or same-direction-sideswipe crashes were less likely to experience severe injuries, whereas those involved in head-on or head-on, left-turn collisions were more likely to experience severe injuries. These results are consistent with past studies and are related to the speed differential and crash force characteristics associated with each crash type (36). Riders who crashed at a stop-controlled or signalized intersection were less likely to experience severe injuries. This finding may be related to the speed of travel when the crash occurred; motorcyclists stopped or moving slowly may be less likely to be injured than those traveling at full speed, especially on a freeway.

Table 1. Michigan Crash Severity Analysis: Summary Statistics

Factor	Number Observed	Percentage of Total	Factor	Number Observed	Percentage of Total
Driver motorcycle endorsement			**Month**		
Endorsed	3,118	49.5	April	247	4
Not endorsed	2,842	45.1	May	953	15.5
Unknown	343	5.4	June	1,184	19.3
Driver's license (state)			July	1,147	18.7
Michigan	5,807	92.1	August	1,146	18.6
Other	283	4.5	September	749	12.2
Unknown	213	3.4	October	502	8.2
Driver helmet use			November	186	3
Yes	4,908	77.9	December	31	0.5
No	838	13.3	**Day of the week**		
Unknown	557	8.8	Weekday	4,030	65.5
Driver age (years)			Weekend	2,115	34.4
<16	22	0.3	**Time of day**		
16 to 29	1,493	23.7	Midnight-3 a.m.	313	5.1
30 to 59	3,608	57.2	3-6 a.m.	215	3.5
≥60	986	15.6	6-9 a.m.	380	6.2
Unknown	194	3.1	9 a.m.-noon	548	8.9
Driver gender			Noon-3 p.m.	1,133	18.4
Male	5,783	91.7	3-6 p.m.	1,596	26
Female	388	6.1	6-9 p.m.	1,185	19.3
Unknown	132	2.4	9 p.m.-midnight	771	12.5
Driver injury severity			Unknown	4	0.1
Fatal (K)	203	3.2	**Weather**		
Incapacitating (A)	1,000	15.9	Clear	4,966	80.8
Nonincapacitating (B)	2,069	32.8	Cloudy	925	15.1
Possible (C)	1,459	23.1	Other	254	4.1
None (O)	1,419	22.5	**Light**		
Unknown	153	2.4	Daylight	4,425	72
Driver impairment			Dark lighted	616	10
Drugs	16	0.3	Dark unlighted	756	12.3
Alcohol	395	6.3	Other	175	2.8
Both	6	0.1	Unknown	3	0.1
Neither	5,886	93.4	**Road condition**		
Passenger helmet use			Dry	5,690	92.6
Yes	516	79.9	Wet	300	4.9
No	81	12.5	Other	152	2.5
Unknown	49	7.6	**Speed limit (mph)**		
Passenger gender			<30	802	13.1
Male	55	8.5	30-50	2,743	44.6
Female	580	89.8	>50	2,544	41.4
Unknown	11	1.7	Unknown	56	0.9
Passenger injury severity			**Crash type**		
Fatal (K)	19	2.9	Single motor vehicle	3,154	51.3
Incapacitating (A)	149	23.1	Head-on	374	6.1
Nonincapacitating (B)	257	39.8	Angle	828	13.5
Possible (C)	123	19	Rear-end	992	16.1
None (O)	91	14.1	Sideswipe	464	7.6
Unknown	7	1.1	Unknown	333	5.4

Table 2. Random Effects Ordered Probit Model: Parameter Estimates

Parameter	**Estimate**	**SE**	**t-Statistic**	**p-Value**
Constant	1.479	0.098	15.1	<0.001
Helmet use	-0.298	0.054	-5.52	<0.001
Age 21-30 years	0.146	0.055	2.65	0.008
Age 51-70 years	0.131	0.048	2.71	0.007
Female	0.322	0.057	5.65	<0.001
Alcohol use	0.547	0.086	6.32	<0.001
Drug use	0.893	0.356	2.51	0.012
Deer involved collision	-0.53	0.078	-6.81	<0.001
Large truck involved collision	0.822	0.191	4.32	<0.001
Single-vehicle collision	-0.318	0.07	-4.52	<0.001
Rear-end collision	-0.857	0.08	-10.73	<0.001
Same-direction sideswipe collision	-0.654	0.101	-6.51	<0.001
Head-on collision	0.369	0.149	2.48	0.013
Head-on or left-turn collision	0.519	0.103	5.04	<0.001
Stop controlled intersection	-0.219	0.075	-2.92	0.004
Signalized intersection	-0.408	0.068	-6.02	<0.001
Horizontal curve on nonfreeway	0.16	0.065	2.45	0.014
Speed limit 40-45 mph	0.25	0.059	4.22	<0.001
Speed limit 50-55 mph	0.218	0.065	3.37	0.001
Speed limit >55 mph	0.297	0.067	4.41	<0.001
Rain or snow	-0.306	0.123	-2.48	0.013
November or December	-0.273	0.123	-2.21	0.027
Weekend	0.104	0.044	2.35	0.019
Thresholds				
Mu(01)	1.021	0.041	25.21	<0.001
Mu(02)	2.527	0.087	28.94	<0.001
Mu(03)	4.128	0.145	28.5	<0.001
SD of random effect				
Sigma	1.090	0.070	15.5600	<0.001

Table 3. Random Effects Ordered Probit Model: Elasticities

	Percentage of Change in Probability of Injury Outcome				
Variable	**O**	**C**	**B**	**A**	**K**
Helmet use	27.2	10.7	-5.4	-26.1	-51.6
Age 21-30 years	-13.8	-5	3	12.6	23.5
Age 51-70 years	-12.5	-4.4	2.8	11.2	20.4
Female	-29	-11.6	5.6	28.3	56.8
Alcohol use	-45.6	-21.3	6.8	49.3	112.9
Drug use	-64.4	-37.4	3.1	81.6	237.5
Deer-involved collision	56.9	13.1	-15.1	-40.9	-61.8
Large truck involved collision	-60.9	-34.2	4.2	75.2	208.9
Single-vehicle collision	30.5	10.5	-6.8	-27	-49.2
Rear-end collision	94.8	17.9	-25.8	-63	-92.2
Same-direction sideswipe collision	72.6	13.9	-20	-48.3	-69
Head-on collision	-32	-14	5.5	33	70.6
Head-on or left-turn collision	-43.4	-20.2	6.5	46.8	106.5
Stop controlled intersection	22.2	6.5	-5.5	-17.9	-29.6
Signalized intersection	42.6	11.1	-10.9	-32.5	-51.3
Horizontal curve on nonfreeway	-15	-5.5	3.2	13.9	26.2
Speed limit 40-45 mph	-23.2	-8.7	4.8	21.7	41.4
Speed limit 50-55 mph	-20.8	-7.2	4.6	18.6	33.9
Speed limit >55 mph	-28.4	-9.8	6.3	25.3	46.1
Rain or snow	31.9	8.4	-8.2	-24.3	-38.4
November or December	28.2	7.6	-7.2	-21.9	-35
Weekend	-9.9	-3.4	2.2	8.8	16

Crashes that occurred along horizontal curves on nonfreeway roads tended to result in more severe injuries, consistent with past studies and probably is a result of restricted sight distances associated with curved road segments (20). Higher speed limits also were associated with more severe injury outcomes for motorcyclists. Crashes occurring on roads with speeds greater than 55 mph (freeways) resulted in a 46.1% increase in the likelihood of a fatal injury, a finding that was not surprising. Crashes occurring in the rain or snow tended to result in less severe injuries, consistent with past studies and most likely due to slower travel speeds and more cautious riding in poor weather (36). Similarly, crashes occurring in the months of November and December tended to result in less severe injuries, most likely due to the winter weather conditions Michigan experiences during these months that would result in slower travel speeds and more cautious riding. Finally, crashes occurring on a weekend tended to result in slightly more severe injuries than crashes occurring on a weekday, maybe because of riskier driving behavior during weekend riding than in daily commuting during the standard work week.

Conclusions

This study adds evidence to inform the continuing debate about the efficacy of UHLs. On April 13, 2012, the State of Michigan repealed its UHL in lieu of a PHL, which requires helmet use only for inexperienced and uninsured riders. After the PHL was enacted, helmet use rates decreased from more than 94% to approximately 73% (35). To understand the broader impacts of this helmet use policy, a detailed, disaggregate-level study was conducted to assess the degree of injury severity sustained by crash-involved motorcyclists before and after Michigan's transition from a UHL to a PHL. By controlling for various rider, roadway, traffic, and weather characteristics, the results showed helmets to reduce the probability of fatalities by more than 50%. Injuries also tended to be less severe in crashes that occurred at intersections, at low speeds, and under inclement weather conditions. Conversely, injuries were more severe in high-speed collisions or when drugs and alcohol were involved in the crash. Female riders, as well as younger (age 21 to 30 years) and older (age 51 to 70 years) riders tended to be more susceptible to injury.

Ultimately, the study results provide additional support for UHLs. Detractors of UHLs often posit that riders should be free to choose whether to wear a helmet and that safety advocates should focus instead on furthering education to encourage helmet use without mandating it by law. Despite evidence to support UHLs, a 2012 survey conducted in Florida found the vast majority of riders in favor of the state's mandatory training law but less supportive of a mandatory helmet law (37).

However, sound evidence has been presented to argue that UHLs are necessary to protect individuals from their own poor choices (38). A recent analysis found that although other measures could lead to increased helmet use rates and fewer injuries, UHLs are the most effective measures for promoting safety and mitigating the economic impacts of injuries and fatalities (39). These findings are echoed by a 2010 study that suggests legislation may be a more effective and efficient means to increase helmet use than educational programs (40).

Helmet use rates in states with UHLs are around 94%, and compliance rates in states without UHLs are around 50% (41). Coupled with results from the overwhelming body of evidence from the research literature and compelling evidence on the effectiveness of helmets demonstrated in this study, this fact suggests that states should carefully consider moving toward UHLs.

References

1. Motorcycle and Bicycle Helmet Use Laws. Insurance Institute for Highway Safety, Arlington, Va. http://www.iihs.org/laws/helmetusecurrent.aspx. Accessed June 3, 2013.

2. Johnson, S., J. Walker, and D. Utter. Research Note: Crash Outcome Data Evaluation System (CODES) Project Safety Belt and Helmet Analyses. National Center for Statistics and Analysis, NHTSA, U.S. Department of Transportation, Feb. 15, 1996.

3. Hooten, K. G., and G.J.A. Murad. Helmeted vs. Nonhelmeted: A Retrospective Review of Outcomes from 2-Wheeled Vehicle Accidents at a Level 1 Trauma Center. Clinical Neurosurgery, Vol. 59, 2012, pp. 126–130.

4. Houston, D.J., and L.E. Richardson. Motorcyclist Fatality Rates and Mandatory Helmet-Use Laws. Accident Analysis and Prevention, Vol. 40, 2008, pp. 200–208.

5. Morris, C.C. Generalized Linear Regression Analysis of Association of Universal Helmet Laws with Motorcyclist Fatality Rates. Accident Analysis and Prevention, Vol. 38, 2006, pp. 142–147.

6. Traffic Safety Facts: Laws. Motorcycle Helmet Use Laws. DOT HS 810 726W. National Center for Statistics and Analysis, NHTSA, U.S. Department of Transportation, February 2007.

7. Rutledge, R., and J. Stutts. The Association of Helmet Use with the Outcome of Motorcycle Crash Injury When Controlling for Crash/ Injury Severity. Accident Analysis and Prevention, Vol. 35, No. 3, 1993, pp. 347–353.

8. Gabella, B., K.L. Reiner, R.E. Hoffman, M. Cook, and L. Stallones. Relationship of Helmet Use and Head Injuries Among Motorcycle Crash Victims in El Paso County, Colorado, 1989–1990. Accident Analysis and Prevention, Vol. 27, No. 3, 1995, pp. 363–369.

9. Bledsoe, G.H., and G. Li. Trends in Arkansas Motorcycle Trauma After Helmet Law Repeal. Southern Medical Journal, Vol. 98, No. 4, 2005, pp. 436–440.

10. Liu, B.C., R. Ivers, R. Norton, S. Boufous, S. Blows, and S. K. Lo. Helmets for Preventing Injury in Motorcycle Riders (Review). Cochrane Database of Systematic Reviews, Vol. 1, 2008.

11. Muller, A. Florida's Motorcycle Helmet Law Repeal and Fatality Rates. Research and Practice, Vol. 94, No. 4, 2004, pp. 556–558.

12. Deutermann, W.V. Traffic Safety Facts: Research Note. Calculating Lives Saved by Motorcycle Helmets. DOT HS 809 861. National Center for Statistics and Analysis, NHTSA, U.S. Department of Transportation, August 2005.

13. Branas, C.C., and M.M. Knudson. Helmet Laws and Motorcycle Rider Death Rates. Accident Analysis and Prevention, Vol. 33, 2001, pp. 641–648.

14. Cummings, P., F.P. Rivara, C.M. Olson, and K.M. Smith. Changes in Traffic Crash Mortality Rates Attributed to Use of Alcohol, or Lack of a Seat Belt, Air Bag, Motorcycle Helmet, or Bicycle Helmet, United States, 1982–2001. Injury Prevention, Vol. 12, 2006, pp. 148–154.

15. Houston, D.J. Are Helmet Laws Protecting Young Motorcyclists? Journal of Safety Research, Vol. 38, 2007, pp. 329–336.

16. Mayrose, J. The Effects of a Mandatory Motorcycle Helmet Law on Helmet Use and Injury Patterns Among Motorcyclist Fatalities. Journal of Safety Research, Vol. 39, 2008, pp. 429–432.

17. Mertz, K.J., and H.B. Weiss. Changes in Motorcycle-Related Head Injury Deaths, Hospitalizations, and Hospital Charges Following Repeal of Pennsylvania's Mandatory Motorcycle Helmet Law. Research and Practice, Vol. 98, No. 8, 2008, pp. 1464–1467.

18. Houston, D.J. The Case for Universal Motorcycle Helmet Laws. Southern Medical Journal, Vol. 103, No. 1, 2010, pp. 1–2.

19. Weiss, H., Y. Agimi, and C. Steiner. Youth Motorcycle-Related Brain Injury by State Helmet Law Type: United States, 2005–2007. Pediatrics, Vol. 126, No. 6, 2010, pp. 1149–1155.

20. Eustace, D., V.K. Indupuru, and P. Hovey. Identification of Risk Factors Associated with Motorcycle-Related Fatalities in Ohio. Journal of Transportation Engineering, Vol. 137, 2011, pp. 474–480.

21. French, M.T., G. Gumus, and J.F. Homer. Motorcycle Fatalities Among Out-of-State Riders and the Role of Universal Helmet Laws. Social Science and Medicine, Vol. 75, 2012, pp. 1855–1863.

22. Fagnant, D.J., B.G. Nichols, and K.M. Kockelman. Who Rides and Who Pays: Comprehensive Assessment of the Costs and Benefits of Motorcycling in the United States. In Transportation Research Record: Journal of the Transportation Research Board, No. 2388, Transportation Research Board of the National Academies, Washington, D.C., 2013, pp. 37–46.

23. Stolzenberg, L., and S.J. D'Alessio. Born to Be Wild: The Effect of the Repeal of Florida's Mandatory Motorcycle Helmet-Use Law on Serious Injury and Fatality Rates. Evaluation Review, Vol. 27, No. 2, 2003, pp. 131–150.

24. The Safety and Economic Costs of Repealing Michigan's Helmet Law. Transportation Research Institute, Ann Arbor, Mich., March 2012. http:// www.smarterusa.org/documents/UMTRI-Mar2012-repeal-analysis.pdf. Accessed June 3, 2013.

25. Karlson, T.A., and C.A. Quade. Head Injuries Associated with Motorcycle Use—Wisconsin, 1991. Morbidity and Mortality Weekly Report, Vol. 43, No. 23, 1994, pp. 423–426.

26. IIHS. More Deaths Follow Weakening of Florida's Motorcycle Helmet Law. Status Report, Vol. 40, No. 8, 2005, pp. 6–7.

27. McKnight, A.J., and A.S. McKnight. The Effects of Motorcycle Helmets upon Seeing and Hearing. Accident Analysis and Prevention, Vol. 27, No. 24, 1994, pp. 493–501.

28. Research Note: Further Analysis of Motorcycle Helmet Effectiveness Using CODES Linked Data. National Center for Statistics and Analysis, NHTSA, U.S. Department of Transportation, January 1998.

29. Eastridge, B.J., S. Shafi, J.P. Minei, D. Culica, C. McConnel, and L. Gentilello. Economic Impact of Motorcycle Helmets: From Impact to Discharge. The Journal of Trauma, Injury, Infection, and Critical Care, Vol. 60, No. 5, 2006, pp. 978–984.

30. Brown, C.V.R., K. Hejl, E. Bui, G. Tips, and B. Coopwood. Risk Factors for Riding and Crashing a Motorcycle Unhelmeted. Journal of Emergency Medicine, Vol. 41, No. 4, 2011, pp. 441–446.

31. Donnelly, E. F. Motorcycle Injuries in Rhode Island. Medicine and Health/Rhode Island, Vol. 92, No. 5, 2009, pp. 185–186.

32. Naumann, R.B., and R.A. Shults. Helmet Use Among Motorcyclists Who Died in Crashes and Economic Cost Savings Associated with State Motorcycle Helmet Laws—United States, 2008–2010. Morbidity and Mortality Weekly Report, Vol. 61, No. 23, 2012, pp. 425–437.

33. Savolainen, P., F. Mannering, D. Lord, and M. Quddus. The Statistical Analysis of Crash-Injury Severities: A Review and Assessment of Methodological Alternatives. Accident Analysis and Prevention, Vol. 43, No. 5, 2011, pp. 1666–1676.

34. Washington, S., M. Karlaftis, and F. Mannering. Statistical and Econometric Methods for Transportation Data Analysis, 2nd ed. Chapman and Hall/CRC, Boca Raton, Fla., 2011.

35. Savolainen, P.T., T. Gates, J.A. Morden, and E. Hacker. 2013 Direct Observation Survey of Motorcycle Helmet Use. Michigan Office of Highway Safety Planning, Lansing, Mich., 2013.

36. Savolainen, P.T., and F. Mannering. Probabilistic Models of Motorcyclists' Injury Severities in Single- and Multi-Vehicle Crashes. Accident Analysis and Prevention, Vol. 39, No. 5, 2007, pp. 955–963.

37. Lee, C., J. Pino, and K. Choi. Lessons Learned from Motorcyclist Surveys: Riders' Attitudes and Behaviors in Florida. In Transportation Research Record: Journal of the Transportation Research Board, No. 2388, Transportation Research Board of the National Academies, Washington, D.C., 2013, pp. 71–78.

38. Jones, M.M., and R. Bayer. Paternalism and Its Discontents. American Journal of Public Health, Vol. 97, No. 2, 2007, pp. 208–217.

39. Derrick, A.J., and L.D. Faucher. Motorcycle Helmets and Rider Safety: A Legislative Crisis. Journal of Public Health Policy, Vol. 30, No. 2, 2009, pp. 226–242.

40. Ranney, M. L., M.J. Mello, J.B. Baird, P.R. Chai, and M.A. Clark. Correlates of Motorcycle Helmet Use Among Recent Graduates of a Motorcycle Training Course. Accident Analysis and Prevention, Vol. 42, 2010, pp. 2057–2062.

41. McCartt, A.T., L. Blanar, E.R. Teoh, and L.M. Strouse. Overview of Motorcycling in the United States: A National Telephone Survey. Journal of Safety Research, Vol. 42, 2011, pp. 177–184.