2017

Bayesian Approach on Quantifying the Safety Effects of Pedestrian Countdown Signals to Drivers

Angela E. Kitali
University of North Florida

BAYESIAN APPROACH ON QUANTIFYING THE SAFETY EFFECTS OF PEDESTRIAN COUNTDOWN SIGNALS TO DRIVERS

By

Angela Edes Kitali

A thesis submitted to the School of Engineering

In partial fulfillment of the requirements for the degree of

Masters of Science in Civil Engineering

UNIVERSITY OF NORTH FLORIDA

COLLEGE OF COMPUTING, ENGINEERING, AND CONSTRUCTION

April 2017

The thesis "Bayesian Approach on Quantifying the Safety Effects of Pedestrian Countdown Signals to Drivers" submitted by Angela E. Kitali in partial fulfillment of the requirements for the degree of Master of Science in Civil Engineering has been

Approved by the thesis committee: **Date**

Dr. Thobias Sando

Thesis Advisor and Committee Chairperson

Dr. Donna Mohr

Dr. Don Resio

Dr. O. Patrick Kreidl

Accepted for the School of Engineering:

Dr. Murat Tiryakioglu

Director of the School

Accepted for the College of Computing, Engineering, and Construction:

Dr. Mark Tumeo

Dean of the College

Accepted for the University:

Dr. John Kantner

Dean of the Graduate School

ACKNOWLEDGEMENTS

Foremost, I would wish to thank the Almighty God for His endless blessings to my life. I would like to convey my immense gratitude to my supervisor, Dr. Thobias Sando, who has generously shared his contagious passion for research. I am grateful for the example he has set by his leadership, expertise, understanding, patience, and dedication, all of which molded me into the engineer and researcher I am present. I would like to thank my committee members, Dr. Donna Mohr, Dr. Don Resio, and Dr. O. Patrick Kreidl, who have never hesitated to provide insight and guidance throughout this journey. I am grateful for the collaborative and supportive work to everyone who assisted me in my thesis.

This project was funded by the United States Department of Transportation (USDOT) grant DTRT13-G-UTC42, and administered by the Center for Accessibility and Safety for an Aging Population (ASAP) at the Florida State University (FSU), Florida A&M University (FAMU), and University of North Florida (UNF). My sincere gratitude also extends to the Florida Department of Transportation (FDOT) and the cities of Jacksonville and Gainesville for providing the roadway data. The opinions, results, and findings expressed in this thesis are those of the author and do not necessarily represent the opinions of the USDOT, the FDOT, the ASAP, the FSU, the FAMU, or the UNF. I am also thankful for funding from the American Association of State Highway and Transportation Officials (AASHTO) scholarship and the UNF Graduate School that has supported my research endeavors, school tuition, and allowed me to travel, present, and participate in different conferences.

Special thanks go to the members of the UNF Transportation lab, for their endless support both in research and social life.

I would wish to show my appreciation to my family and friends who have loved, encouraged, and cheered me on along the path. Your support is priceless to where I am today.

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

Page

LIST OF ACRONYMS

AADT	Annual average daily traffic
BCI	Bayesian credible interval
CI	Confidence interval
CMF	Crash modification factor
CMFunction	Crash modification function
CRF	Crash reduction factor
EB	Empirical Bayes
FB	Full Bayes
FI	Fatal and injury
HMC	Hamiltonian Markov chain
MCMC	Markov chain Monte Carlo
MUTCD	Manual on Uniform Traffic Control Devices
PCS	Pedestrian countdown signal
PDO	Property damage only
SPF	Safety performance function

ABSTRACT

Pedestrian countdown signals (PCSs) are viable traffic control devices that assist pedestrians in crossing intersections safely. Despite the fact that PCSs are meant for pedestrians, they also have an impact on drivers' behavior at intersections. This study focuses on the evaluation of the safety effectiveness of PCSs to drivers in the cities of Jacksonville and Gainesville, Florida. The study employs two Bayesian approaches, before-and-after empirical Bayes (EB) and full Bayes (FB) with a comparison group, to quantify the safety impacts of PCSs to drivers. Specifically, crash modification factors (CMFs), which are estimated using the aforementioned two methods, were used to evaluate the safety effects of PCSs to drivers. Apart from establishing CMFs, crash modification functions (CMFunctions) were also developed to observe the relationship between CMFs and traffic volume.

The CMFs were established for distinctive categories of crashes based on crash type (rear-end and angle collisions) and severity level (total, fatal and injury (FI), and property damage only (PDO) collisions). The CMFs findings, using the EB approach indicated that installing PCSs result in a significant improvement of driver's safety, at a 95% confidence interval (CI), by a 8.8% reduction in total crashes, a 8.0% reduction in rear-end crashes, and a 7.1% reduction in PDO crashes. In addition, FI crashes and angle crashes were observed to be reduced by 4.8%, whereas a 4.6% reduction in angle crashes was observed. In the case of the FB approach, PCSs were observed to be effective and significant, at a 95% Bayesian credible interval (BCI), for a total (Mean = 0.894, 95% BCI (0.828, 0.911)), PDO (Mean = 0.908, 95% BCI (0.838, 0.953)), and rear-end (Mean = 0.920, 95% BCI (0.842, 0.942)) crashes. The results of two crash categories such as FI (Mean = 0.957, 95% BCI (0.886, 1. 020)) and angle (Mean = 0.969, 95% BCI (0.931, 1.022)) crashes are less than one but are not significant at the 95 % BCI.

Also, discussed in this study are the CMFunctions, showing the relationship between the developed CMFs and total entering traffic volume, obtained by combining the total traffic on the major and the minor approaches. In addition, the CMFunctions developed using the FB indicated the relationship between the estimated CMFs with the post-treatment year. The CMFunctions developed in this study clearly show that the treatment effectiveness varies considerably with post-treatment time and traffic volume. Moreover, using the FB methodology, the results suggest the treatment effectiveness increased over time in the post-treatment years for the crash categories with two important indicators of effectiveness, i.e., total and PDO, and rear-end crashes. Nevertheless, the treatment effectiveness on rear-end crashes is observed to decline with post-treatment time, although the base value is still less than one for all the three years. In summary, the results suggest the usefulness of PCSs for drivers.

CHAPTER 1: INTRODUCTION

Overview of Pedestrian Countdown Signals (PCSs)

Enhancing pedestrian safety at signalized intersections, especially in urban areas has always been a challenge to traffic engineers. Signalized intersections in urban environments are characterized with high traffic volume and fast approach speeds when crossing intersections. This situation imparts a challenge to pedestrians who intend to cross these intersections. To address this challenge, a number of pedestrian controls have been introduced to improve the safety of pedestrians as they cross signalized intersections. Pedestrian countdown signals (PCSs) are one of the recent pedestrian control devices that have been introduced for the purpose of improving pedestrian safety at signalized intersections. These are standard devices that inform pedestrians that is safe to cross intersections when there is no conflicting vehicular traffic.

A unique feature of PCSs is the display of the remaining time in seconds in the protected pedestrian phase, providing pedestrians with a chance to decide whether they will have enough time to clear the intersection before the termination of their right-of-way. In addition, for pedestrians who have started to clear the intersection, the PCS' timer provides them with an opportunity to speed up to walk through the intersection. It is worth noting that PCSs are replacing the traditional pedestrian signals (Figure 1.1, left picture).

A standard pedestrian signal consists of a steady WALKING PERSON, signifying WALK, which gives permission to a pedestrian to start crossing the intersection in the direction of the signal indication. This pedestrian signal also consists of a flashing and steady UPRAISED HAND, implying DON'T WALK, indicating that a pedestrian shall not start to cross the roadway in the direction of the signal indication, but that any pedestrian who has already started to cross on a steady WALKING PERSON (symbolizing WALK) signal indication shall speed up to clear the intersection. Lastly, a standard pedestrian signal also consists of a steady UPRAISED HAND, symbolizing DON'T WALK, meaning that a pedestrian shall not enter the roadway in the direction of the signal indication (MUTCD, 2009).

Pedestrians have a tendency to be confused when the standard pedestrian signal shows a flashing upraised hand. Though this symbol is meant to provide pedestrians who have already started crossing the roadway sufficient clearance time to finish the crossing safely, some pedestrians and other road users have difficulty estimating the actual amount of time indicated during the flashing hand phase. Presumably, they consider the amount of time designated for the flashing hand phase to be insufficient for pedestrians to finish crossing the intersections safely. This confusion has been solved by the display of the countdown timer in a PCS (Figure 1, right). PCSs were first approved and incorporated in the Manual of Uniform Traffic Control Devices (MUTCD) in its 2003 version (MUTCD, 2003). In 2009, the MUTCD warranted the installation of PCSs at each intersection with a pedestrian clearance interval of more than 7 seconds (MUTCD, 2009). Since then, they have been widely utilized by transportation authorities as a preferred pedestrian treatment at signalized intersections.

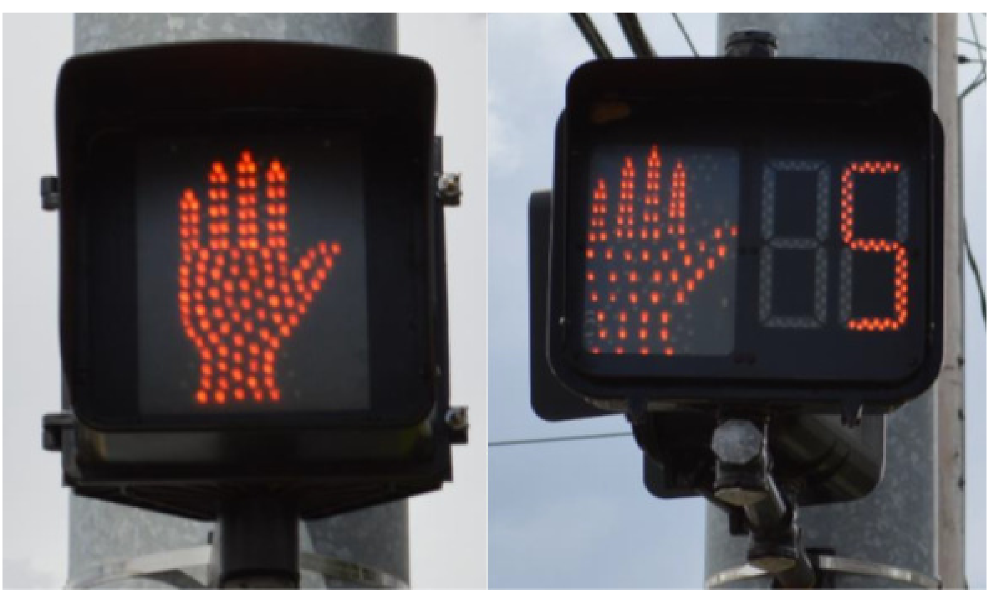


Figure 1.1. Standard pedestrian signal (left) vs pedestrian countdown signal (right)

PCSs have a countdown timer, which exhibits the amount of time remaining in seconds for pedestrians to clear the intersection before their time for crossing is terminated. The remaining time displayed in these signal heads is in descending order, in seconds, using Arabic numerals. Normally PCSs are installed at a distance that ranges between 1.5 and 6 feet away from the curb face (MUTCD, 2009). Further, the PCSs' heads are required to be installed at a height that ranges between 7 and 10 feet above the sidewalk level (MUTCD, 2009). Based on the location and height of PCSs at signalized intersections, some drivers are able to see the countdown timing displayed as they approach intersections on a lane parallel to the pedestrian crosswalk where the PCS is operating.

Often, the operation of PCSs is simultaneous to the commencement of the green phase on the approaching through traffic parallel to the pedestrian walkway where the PCS is directing pedestrians to cross. The time displayed on the operating PCS' head is also visible to motorists on approaches parallel to the crosswalk. When the pedestrian clearance interval reads zero, the approaching traffic parallel to the pedestrian walkway will commonly receive a red signal, with the exception of some of the intersections where the actuated green signal may be extended depending on the detected length of the vehicle platoon waiting to cross the intersection. Thus, the information that a pedestrian countdown timer offers to approaching drivers may affect their approach speed, especially when drivers realize that the termination of the PCS' timer operation is simultaneous to the termination of their green phase.

Effects of PCSs on Drivers

There are two schools of thought regarding the drivers' reactions as the pedestrian countdown timers near zero. There are drivers who would use the timer as a cue to speed up to clear the intersection, avoiding being stopped by the traffic light changing from green to red. On the other hand, when the timer approaches zero, some drivers slow down and prepare to stop.

Considering the fact that PCSs are installed with the main purpose of improving the safety of pedestrians when crossing intersections, many studies have focused on evaluating pedestrian behavior responses towards PCSs (Kimberly et al., 2004; Scott et al., 2012; Vasudevan et al., 2011). There is an agreement in the research community on the effectiveness of PCSs on the improvement of pedestrian safety and operations as they cross signalized intersections. However, research on the impacts that PCSs have on drivers is more limited.

A study by Schmitz (2011) observed the speeds of vehicles at the stop bar decreased by 1.0 mph in the presence of PCSs as compared to intersections without PCSs. In contrast, other drivers, upon spotting countdown timing, are encouraged to speed through the intersection to finish crossing the intersection before the termination of the green phase. In a study conducted to evaluate the influence of PCSs on vehicle speeds as they approach the intersection (Nambisan and Karkee, 2010), results showed that when a PCS' timer is counting down vehicle speeds tend to be higher near the intersections than the roadway segment upstream of the intersections. This indicates that drivers utilize information displayed on PCS' timers to speed up and clear the intersection to avoid being caught waiting for the next cycle. The results of this study also indicated that speeds of vehicles approaching the intersection are not influenced by the actual numeric display on the PCS' timer. Another study by Chen et al. (2015) observed more red light violation and early-start maneuvers at signalized intersections with PCSs for both motorcyclists and drivers compared to intersections without PCSs. While studies on the effects of PCSs on pedestrians have consistently reported improvements in pedestrian safety, research on the safety effectiveness of PCSs for drivers appear to yield conflicting findings. A study is therefore warranted to appraise the safety effects of these signals to drivers.

Measures to Evaluate the Safety Effects of PCSs to Drivers

Most transportation agencies promote the use of crash modification factors (CMFs) for evaluating safety effectiveness of improvements made to roadway facilities. In fact, CMFs for countdown signals are included in the CMFs' most-wanted list on the CMF clearinghouse website (HSRC, 2016). A CMF is an amplification factor employed to quantify the safety effectiveness of the installed countermeasure at a specific site. Implemented treatments include variations in the geometric and traffic characteristics of roadway facilities. Varies in safety changes is measured relative to a reference value given a CMF of 1.0, which indicates that presence of the treatment did not influence changes in the crash frequency. For example, for an improved treatment such as intersections with PCSs, a CMF of 0.98 represents an anticipated two percent (2%) reduction in crashes. On the other hand, the CMF of 1.02 indicates an anticipated deterioration in safety, i.e. a two percent (2%) increase in crashes. In order to capture the safety impact of the respective treatment, a crash reduction factor (CRF), which is a reverse of CMF (CRF=1-CMF), is computed.

Together CMFs and CRFs are important measures of effectiveness commonly used by transportation professionals for different purposes. Specifically, they assist in the assessment of the safety effects of different installed countermeasures and compare safety impacts among numerous alternatives and locations. Further, CMFs and CRFs can be used to categorize cost-effective strategies and locations based on crash effects (Gross et al., 2010). They are also used for analyzing the economic impact of proposed safety countermeasures (NCHRP, 2008; AASHTO, 2010).

CMFs provide an overall estimate of the safety impacts of a treatment. To estimate CMFs, the expected number of crashes on the treatment intersections during the after period—assuming the treatment was not installed— is computed. Specifically, the CMF is computed as the annual observed crashes on the treatment sites divided by the expected number of crashes on the treatment intersection assuming the absence of the treatment. By this definition, the CMF is synonymous to the odds ratio. Estimation of the expected number of crashes in the absence of the countermeasure is affected by a number of factors, including the regression-to-the-mean (RTM) effect. RTM elaborates the situation in which crash rates are relatively high before the installation of the countermeasure where they tend to come back to the mean in the year following an unexpectedly high or low crash count. In most cases the proposed safety treatment targets high-hazard sites. Therefore, safety analysis of the installed treatments is likely to be affected by the RTM phenomenon simply because the sites experience an immediate reduction in the number of crashes after these counts come back to their average mean (Gross et al., 2010). Figure 1.2 summarizes the annual data for the Hodges Boulevard and Beach Boulevard intersections, located in Jacksonville, retrieved from the Signal Four Analytics crash database.

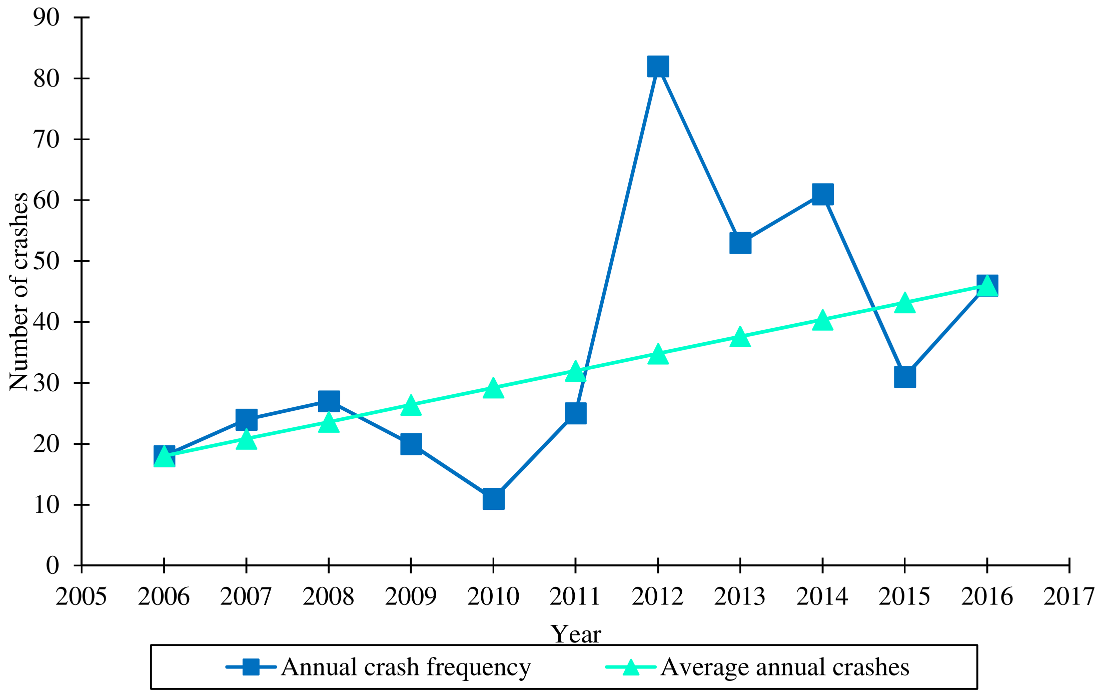


Figure 1.2. Data Series for Hodges Blvd and Beach Blvd Intersection (example intersection)

From Figure 1.2, it can be observed that the trend in crash frequency is increasing. An interesting insight from this plot is that each time a higher number of crashes are observed, they tend to regress to the common moving average. A change in number of crashes would be anticipated eventually even if a treatment was not installed. It is possible to over-estimate the effects of the installed treatment on these sites if the RTM bias is not properly addressed in the analysis (Gross et al., 2010). Apart from the RTM effects, the estimation of the expected crashes on the treatment intersections is expected to be affected by changes in traffic volume, unrelated factors, and trends (Sacchi & Sayed, 2014; Gross et al., 2010). Thus, it is important that the estimated CMF represents the long-term expected change in crash frequency (Gross et al., 2010). The actual change in crashes observed after treatment will vary by location and time (year).

To examine the variation of the estimated CMF across sites and at different post-treatment years, crash modification functions (CMFunctions) are used. A CMFunction is a formula used to estimate the CMF for a specific site based on its heterogeneous characteristics and also within different post-intervention periods (Gross et al., 2010). The CMFunction allows the CMF to vary with the change in one or a combination of variables such as traffic volume, geometric characteristics, and land use information.

Empirical Bayes versus Full Bayes

Various techniques are used to estimate CMFs where the observational before-and-after methods are the most preferred ones (Gross et al., 2010; Sacchi & Sayed, 2014). The superior benefit of the before-and-after methods over other safety effectiveness methods is that it is a longitudinal analysis, i.e., CMFs derived from the before-and-after studies are based on the change in safety performance due to the installed countermeasure. A reliable before-and-after study method should ensure that a change in safety has been influenced by the countermeasure and not by other external confounding attributes (Sacchi & Sayed, 2014). These include change in traffic volume, time trends and unobserved heterogeneity. In this regard, the conventional empirical Bayes (EB) method with a comparison group is one of the safety effectiveness methods which is able to account for the above-mentioned potential biases (Hauer, 1997).

The EB method with a comparison group is a statistical approach that more precisely combines the observed crash frequency with the predicted crash frequency using the safety performance function (SPF). An SPF, derived from the comparison group, is used to calculate the expected crash frequency for treatment sites assuming the countermeasure not been implemented (Figure 1.3).

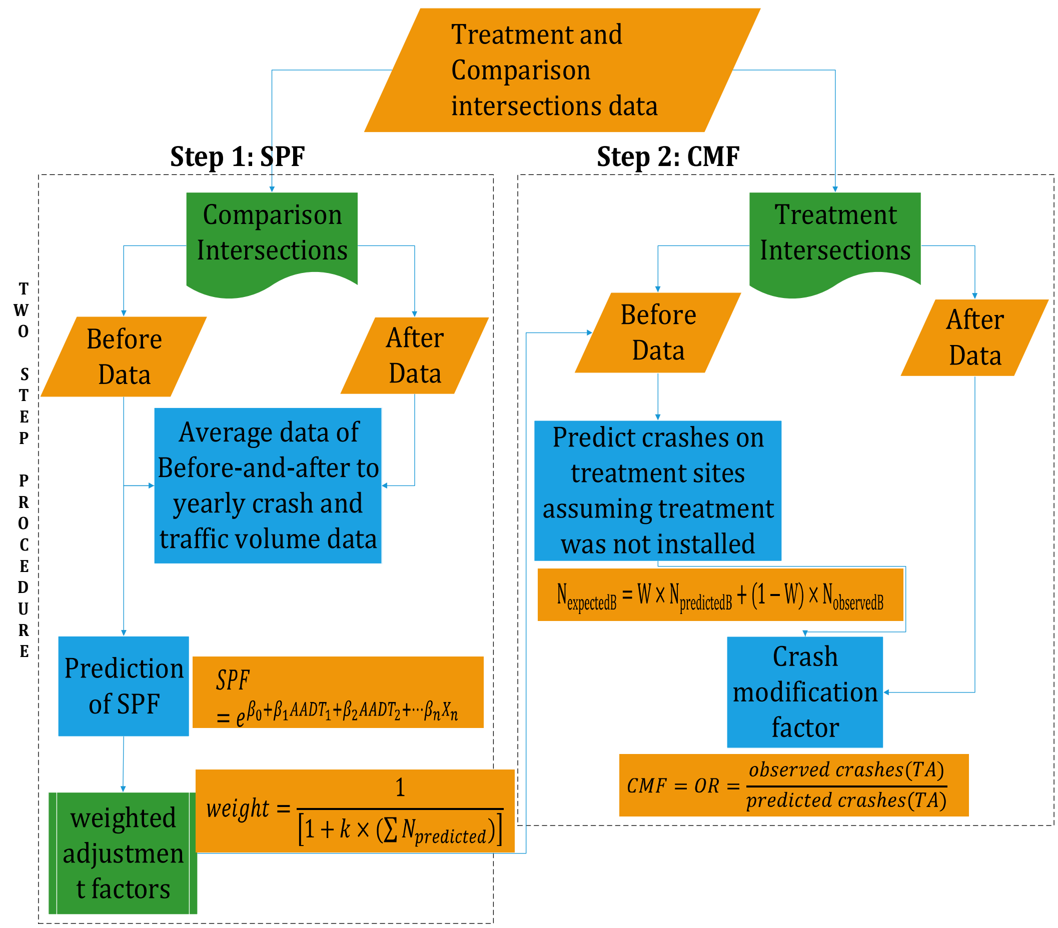


Figure 1.3. Empirical Bayes flow chart

The benefit of the EB method is that it correctly accounts for the observed changes in crash frequencies before-and-after treatment that may be influenced by the RTM effect (Gross et al., 2010). It is worth noting that the comparison group refers to sites with similar geometric, traffic, and land use information to the treatment sites but without treatment during the after- period. The data from comparison sites are utilized to account for changes in crashes unrelated to the treatment, such as time and traffic volume trends (Gross et al., 2010). Specifically, the data from this group are applied to compute the ratio of the observed crash frequency in the after-period to that in the before-period. The estimated ratio is then multiplied by the observed crash frequency in the before-period at a treatment site group to provide estimates of expected crashes on the treatment sites had the countermeasure not been applied (Figure 1.4).

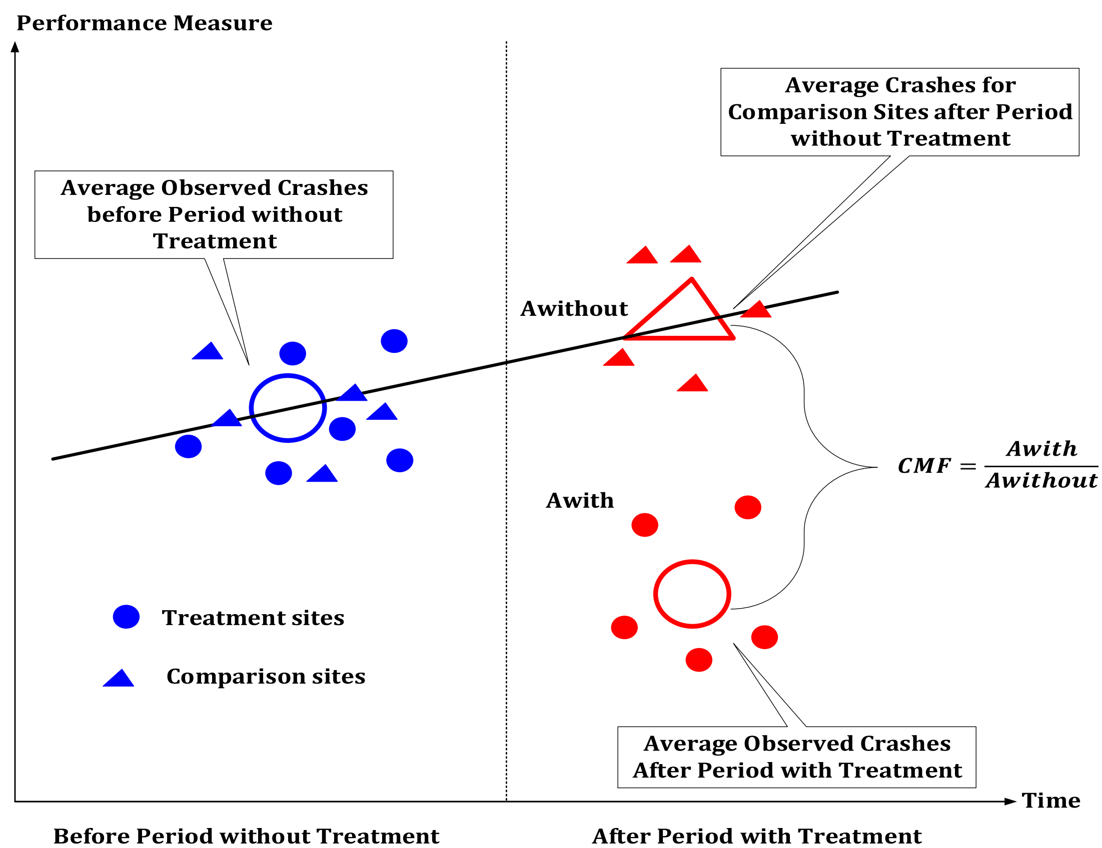


Figure 1.4. Before-and-after evaluation using a comparison group Source: (Herbel et al., 2010)

Alternatively, CMFs can be derived using another Bayesian approach, full Bayes (FB). Both EB and FB methods employ the same concepts to account for external confounding effects in appraising the safety effectiveness of the countermeasures. Nonetheless, there are a number of appealing characteristics of the FB method. The FB approach has the capability to account for most of the uncertainties in the dataset and model parameters and hence overcome the maximum likelihood methods' problem of overestimating precision because of ignoring this uncertainty (Park et al., 2016). This approach has the ability to include prior knowledge on the values of the coefficients in the model along with the observed data (Gross et al., 2010). The FB methodology is also a single-step integrated procedure (Figure 1.5), i.e., it integrates the process of estimating the SPF and the treatment effect in a single step, thus incorporating the uncertainties of the SPFs in the final estimates.

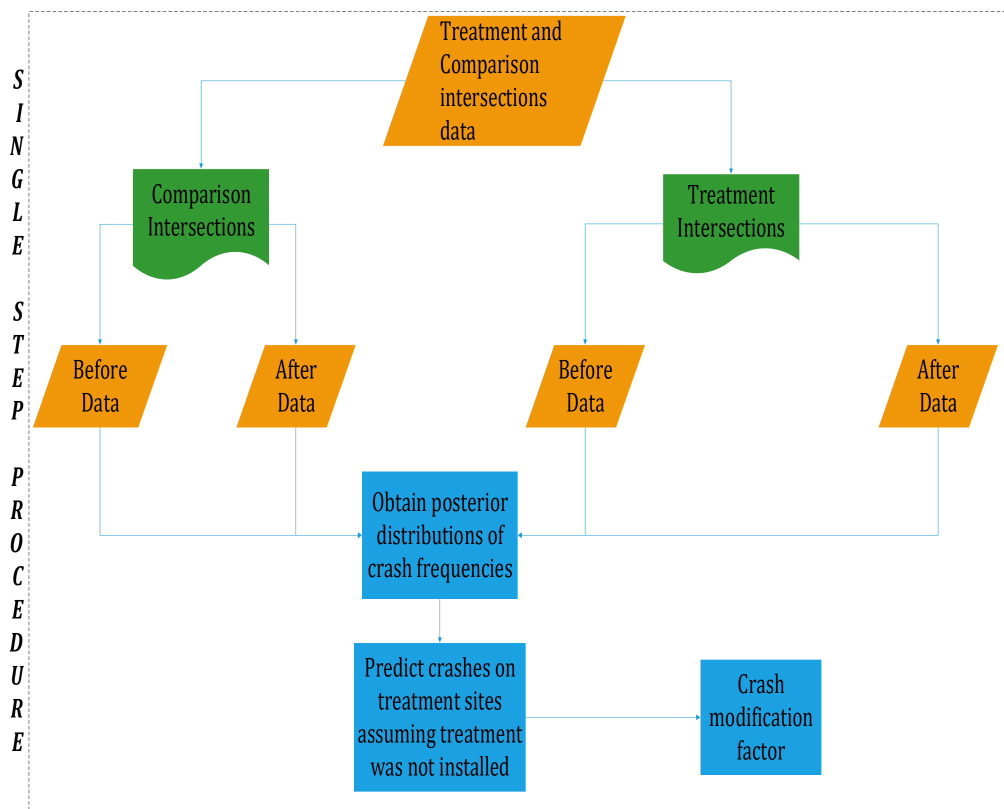


Figure 1.5. Full Bayes flow chart

Another advantage is that the properties of the FB method allow for the computation of reliable models with smaller sample sizes (Li et al., 2013; Ahmed, et al., 2015; Gross et al., 2010). This characteristic may be specifically valuable to rare crash categories such as those involving pedestrians or for comparison cohorts with limited sites, for example five-legged intersections. Other benefits of the FB method include the ability to consider the spatial correlation between sites in the model formulation (Gross et al., 2010).

Study Objectives

The main objective of this study is to examine the influence of PCSs on driver safety. In order to accomplish this objective, this study employed two major approaches used to quantify the safety impacts of an installed treatment. These approaches are EB and FB approaches used in developing crash modification factors and functions.

Thesis organization

This thesis comprises four chapters. Chapter 1 provides the general overview of the research problem and the description of the research objectives. Because this thesis is a compilation of two research articles, chapters 2 and 3 are two stand-alone research papers that are based on the subject at hand. Chapter 4 summarizes, concludes, and provides a list of opportunities for future work based on the limitations of this thesis.

CHAPTER 2: PAPER 1

PAPER I: Appraisal of Safety Effects of Pedestrian Countdown Signals to Drivers Using Crash Modification Factors

Paper I was originally published in the 2017 Transportation Research Board (TRB) annual meeting compendium of papers. The same paper was also presented at the 96th TRB annual meeting in January 2017 in Washington, D.C. A variation of this paper has been submitted and is being considered for publication in the journal of Transportation Engineering (JTE) Part A. Also, Paper 1 received the Research Recognition Award during the SOARS 2017 presentations and selected for presentation at the statewide symposium in April 2017.

Introduction

Pedestrian countdown signals (PCSs) were first approved and included in the 2003 version of the Manual of Uniform Traffic Control Devices (MUTCD). Since then, they have been widely used by different transportation agencies as a preferred pedestrian treatment at signalized intersections. For pedestrians, PCSs provide convenient information that helps them to cross the street safely. Although the safety benefits of PCSs on pedestrian safety are well established by research (Huang and Zeeger, 2000; Markowitz et al., 2006; Chen et al., 2015; Lambrianidou et al., 2013; Schmitz 2011; Scott et al., 2012; Vasudevan et al., 2011; and Eccles et al. 2004), the effects of PCSs on driver's safety are still debatable.

PCSs incorporate countdown timer clocks, which displays the amount of time remaining in seconds for pedestrians to clear the intersection before their crossing time is terminated. The remaining time is displayed in descending order using the Arabic numerals. The MUTCD requires PCSs to be used when the pedestrian change interval is above 7 seconds (MUTCD, 2009). Normally, PCSs are installed at a distance that ranges between 1.5 and 6 feet away from the curb face according to MUTCD (2009). Furthermore, the MUTCD requires

PCS' heads to be installed at a height that ranges between 7 and 10 feet above the sidewalk level (MUTCD, 2009). Based on the location and height of PCSs at signalized intersections, some drivers are able to see the countdown timing displayed as they approach intersections. Hence, the cues that PCS,' timers provide to drivers may impact their safety at signalized intersection. This situation call for a need to evaluate the safety effects of these signals to drivers.

Most transportation agencies promote the use of crash modification factors (CMFs) for evaluating safety effectiveness of improvements made on roadway facilities. In fact, CMFs for countdown signals are included in the CMFs' most-wanted list on the clearinghouse website (HSRC, 2016). A CMF is an amplification factor that represents potential variation in the expected number of crashes following implementation of a specific treatment. Implemented treatments include changes in the geometric and traffic characteristics of roadway facilities. Variation in safety changes are measured relative to a baseline value that is assigned a CMF of 1.0. For example, for an improved treatment such as intersections with PCSs, if the CMF is 0.98 and the comparison site is intersections without PCSs, an intersection with PCSs is expected to experience a 2 percent (2%) reduction in crashes following the installation of PCSs. In order to capture the safety impact of the respective treatment, a crash reduction factor (CRF), which is a reverse of CMF (CRF=1-CMF), is computed. Both CMF and CRF are important measures of effectiveness commonly used for analysis of the costs and benefits of proposed safety countermeasures (NCHRP, 2008). Specifically, they assist in the selection of improvement projects by quantifying the benefits from potential crash reduction associated with each proposed countermeasure (AASHTO, 2010).

The main objective of this study is to appraise the safety impact of installing PCSs, at signalized intersections, to drivers. In order to fill the knowledge gap identified on the CMFs' clearinghouse website, this study develops CMFs for PCSs. In the process, the study develops crash modification functions for total entering traffic through observational before-and-after study using the empirical Bayes method.

Literature Review

Limited research has been done on safety implications of PCSs to drivers at intersections. According to the study that evaluated the engineering improvements of older drivers in Michigan, it was observed that the presence of PCSs at signalized intersections reduces not only pedestrians' crashes, but also they were found to reduce five percent (5%) of total crashes for all drivers (Kwigizile et al. 2015). This indicates that drivers utilize information provided by PCS' timers to make informed decisions, when approaching and crossing signalized intersections (Chen et al., 2015; Schmitz, 2011; Elekwachi, 2010; and Nambisan and Karkee, 2010). While, some of the drivers use the information on the PCS' timer clock to slow down as they approach the intersection prior to termination of their green phase, other drivers use the same information to speed up to clear the intersection. Thus, they avoid being stopped and waiting for the next cycle. Presented henceforth, is the summary of the literature from studies related to driver behaviors toward PCSs.

Elekwachi (2010) conducted an empirical study to investigate the effect of PCSs on driver behaviors and capacity at signalized intersections. PCSs were observed to have a statistically significant impact on driver behaviors and intersection capacity. They were further noticed to improve driving decisions and affects braking or stopping maneuvers at signalized intersections. Another study conducted by Pulugurtha et al. (2010) found that drivers use the information on PCS' timer clock to decide on slowing down prior to the onset of a yellow phase. On the other hand, another study has reported the speed decrease by 1.0 miles per hour (mph) on locations with PCSs compared to locations without PCSs (Schmitz, 2011).

Conflicting results have been reported by researchers, Chen et al. (2015) and Nambisan and Karkee (2010), indicating that some drivers, upon spotting countdown timing, are encouraged to speed through the intersection, in order to clear the intersection before the termination of their green phase. In a research study that evaluated the influence of PCSs on vehicle speeds as they approach the intersection, vehicles were observed to maneuver at higher speeds on the segment closer to the intersection than on the segment farther away from the stop bar (Nambisan and Karkee, 2010). The results of the study by Chen et al. (2015) indicated a prevalence of red-light violation and early-start maneuvers at signalized intersections with PCSs as compared to intersections without PCSs. This behavior was observed to be critical on both drivers and motorcyclists. The proportion of vehicles entering the intersection late in the yellow phase and red-light runners increases because some of the drivers used PCS' timer information to speed up in order to clear the intersection legally. These conflicting scenarios may create different expectations among drivers, the situation that might result in rear-end conflicts. A possible conflict scenario may occur when the leading vehicle stops while the driver of the following vehicle decides to accelerate to clear the intersection, (Park et al., 2016; Long et al., 2013).

Apart from safety impacts, previous research has documented the influence of PCSs in operational characteristics. Elekwachi (2010) has reported the influence of PCSs on intersection operational characteristics such as headway, saturation flow rate, capacity, start-up lost time, and driver behaviors. According to the study, the presence of PCSs reduces headway, hence increasing the saturation flow, in view of the fact that drivers in the queue are aware of the upcoming phase change. The study also found that PCSs significantly reduce the amount of start-up lost time. This may be attributed to the fact that drivers waiting in the queue at an intersection are aware of the number of seconds remaining on the opposite phase. Thus, they respond quicker to the changing phase.

Methodology

Data description

The analysis required two sets of data: crash data and traffic volumes for before-and-after installation of PCSs. It is important to note that, the sites selected for evaluating the safety effectiveness of a treatment have to be homogenous as recommended by the Section C.5 of the Highway Safety Manual (HSM) (AASHTO, 2010). Among the potential characteristics that have been proposed to be used in identifying treatment sites for intersections include traffic control, i.e. signalized, for this case and a number of approaches e.g. four-legged or three-legged intersections. Considering the limit of the number of three-legged signalized intersections at the study area, the study was limited to only four-legged signalized intersections. This is because another vital criteria to consider while collecting sites for performing safety effectiveness studies is sufficiency of sample size such that the expected change in safety can be statistically detected (Gross, 2010). Furthermore, site selection process was limited only to state maintained roadways due to reliability of traffic volume data (Average Annual Daily Traffic (AADT)) for these sites.

One hundred and ten (110) signalized intersections with PCSs in Jacksonville and Gainesville, Florida were selected as treatment sites for this study, where PCSs were installed between years 2006 through 2011. This includes 70 intersections in Jacksonville, and 40 in Gainesville. For each of the treatment intersection, three years before the installation of PCSs and three years after the installation of PCSs were used for analysis of changes in crash frequency due to installation of PCSs. It is worth mentioning that, the respective year that a PCS was installed in each of the intersection was excluded from the study to allow enough buffer time for changes brought about by PCSs.

In the site selection procedure, 93 comparison sites, i.e., intersections without PCSs, 33 in Gainesville, and 60 in Jacksonville, were carefully selected according to their geographic proximity to the treatment sites (intersections with PCSs), while keeping the distance to avoid a spillover effect. In general, the comparison intersections were selected within the area of the same municipality that share similar geometric characteristics, traffic volume and crash frequencies, as the treatment sites. This was so to improve the comparability between comparison and treatment sites.

Data were collected and retrieved from the following databases; Florida Unified Base-Map Repository (crashes), Florida Geographic Data Library (FGDL) Metadata explorer (land use information), and Florida Department of Transportation (FDOT) Geographical Information System (GIS) database (posted speed). Other sources of data included Google Earth-street view and historical imagery tool. These were used to retrieve geometric information from previous year's before-and-after installation of PCSs. The last but not least data source used in this study was the Florida Traffic Monitoring Sites (TMS) where traffic volume for years 2003 through 2014 were obtained. The historical imagery tool in Google Earth Pro software was utilized in ensuring the quality of the developed SPFs by checking the treatment sites to verify that there is no major geometric change during the study period. The flow chart in Figure 2.1 illustrates the strategy employed in collecting data for this study. The solid and dotted lines indicate the process for collecting data for treatment and comparison intersections, respectively.

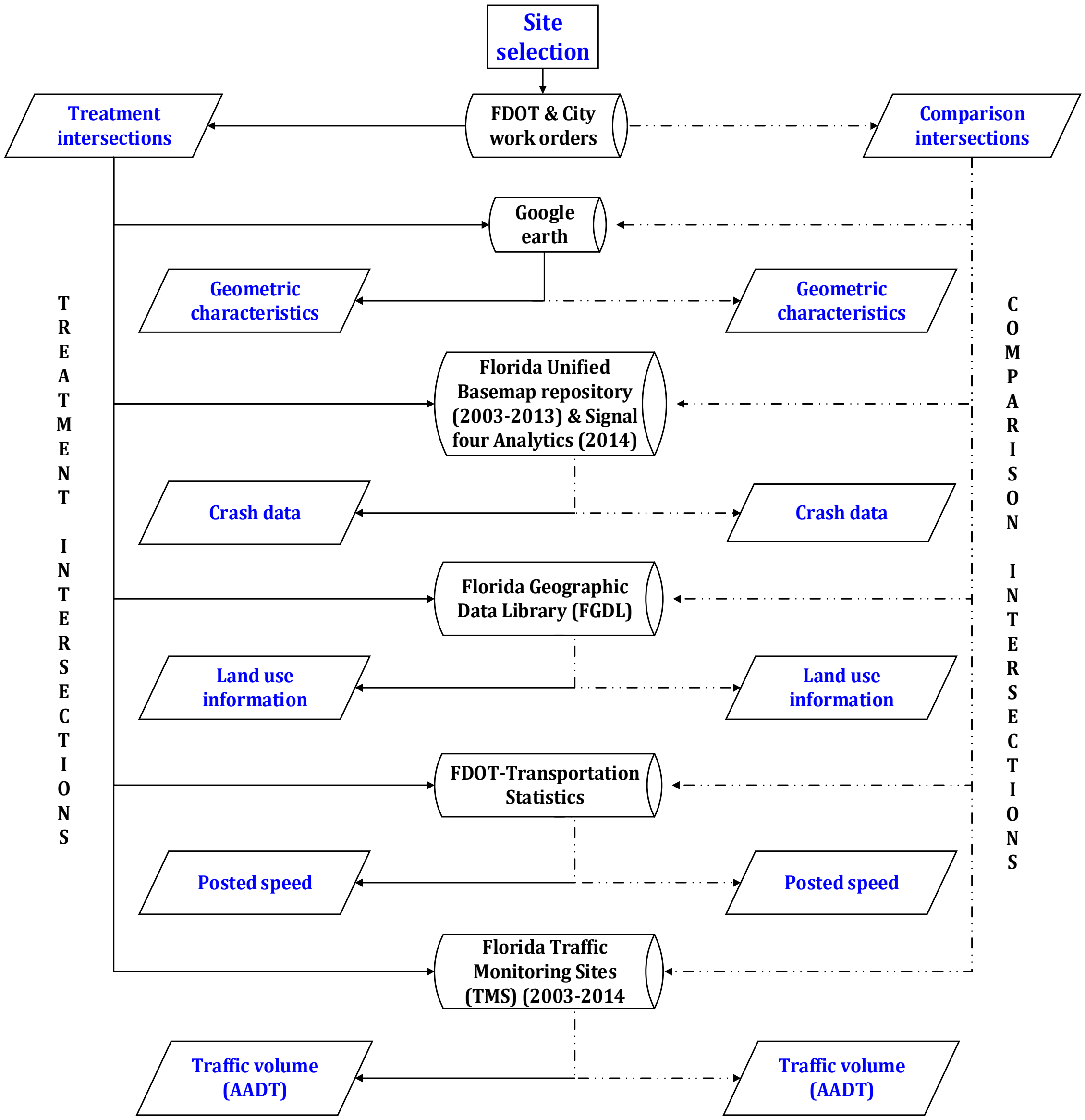


Figure 2.1. Flow chart illustrating data collected in this study

Crash data were available from the years 2003 through 2014. Table 2.1 summarizes the crash data categorized in crash severity (total, fatal-plus-injury (F+I), and property-damage-only (PDO) collisions) and crash types (rear-end and angle collisions). All crashes that occurred within 250 feet were considered to be intersection related. This radius of 250 feet conforms to the definition of intersection related crashes in Florida (FDOT, 2012). In Florida, crashes occurring between 0 and 50 feet from the intersection are referred to as intersection crashes whereas crashes occurring between 50 and 250 feet are termed influenced by intersection. In Florida, when analyzing intersection safety, all crashes that occur within 250 feet radius from the middle of the intersection is considered with the assumption that they are influenced by the presence of the intersection (Jacob, 2015).

Table 2.1 Annual Crash Data Summary for Treatment and Comparison Sites

No.	**Type of crash**	**Period**	**Treatment sites statistics**				**Comparison sites statistics**			
			Mean	SD	Min	Max	Mean	SD	Min	Max
1.	Total crashes	Before	18.38	17.94	0	87	19.97	20.40	0	69
	Total crashes	After	14.31	14.56	0	74	13.23	12.25	1	64
2.	F + I crashes	Before	11.73	11.79	0	58	10.79	9.27	0	40
	F + I crashes	After	10.55	9.56	0	40	11.03	10.22	0	32
3.	PDO crashes	Before	10.50	12.93	0	83	10.91	10.34	0	52
	PDO crashes	After	9.30	10.08	0	60	9.06	11.93	0	68
4.	Rear-end crashes	Before	10.09	12.76	0	62	11.65	10.44	0	58
	Rear-end crashes	After	6.74	8.07	0	44	7.23	8.57	0	46
5.	Angle crashes	Before	5.29	5.47	0	28	5.00	5.09	0	13
	Angle crashes	After	4.75	5.16	0	26	3.78	4.82	0	10

Note: SD= standard deviation, Min=minimum, Max=maximum.

Important variables considered for this study includes traffic volumes on the major and minor approaches, geometric characteristics of major and minor approaches, and the area type (commercial categorized as 1 and other land use types as 0) as illustrated in Table 2.2. In this study, the number of lanes ranged from 2 to 6. If the variable for number of lanes was considered to be categorical, it would have had five categories. Normally, for categorical variables, few categories are associated with more robust model outcome. To avoid causing model instability in relation to the number of lanes category, this study considered number of lanes variable to be continuous as employed by various previous studies (Agbelie & Roshandeh, 2014; Chen et al., 2009; Chen et al., 2016).

Table 2.2 Selected Variables in Treatment and Comparison Intersections

	Variable	Site	Mean	SD	Min	Max
	Land use (1 commercial, otherwise 0)	Treatment	0.47	0.50	0	1
		Comparison	0.70	0.46	0	1
Major	Average AADT (Vehicle/day)	Treatment	28894.20	12281.44	6034	57459
		Comparison	29041.13	12750.75	6400	56750
	Total number of lanes	Treatment	5.33	0.95	4	6
		Comparison	5.78	0.78	4	6
	Posted speed (mph)	Treatment	41.18	6.74	30	55
		Comparison	40.89	5.92	30	55
Minor	Average AADT (Vehicle/day)	Treatment	11427.93	8120.44	567	36000
		Comparison	12839.82	7418.19	840	37500
	Total number of lanes	Treatment	2.93	1.00	2	4
		Comparison	3.01	1.32	2	4

Note: AADT=annual average daily traffic.

Other information retrieved on the major street include, posted speed (>40 mph coded as 1, otherwise 0). It is worth noting that the same cut-off point of the posted speed has been used by Donell et al. (2014) while developing the SPFs. Also, when examining the dataset, it was observed that almost half of the sites used for this study had posted speed limit between 30 mph to 40 mph, and the remaining group (45 mph and above) also constituted about half of the data.

Modeling approach

This study uses the before-and-after empirical Bayes (EB) method illustrated by Hauer (1997), to develop CMFs for PCSs. The advantage of the empirical Bayes approach is that it accounts for the observed changes in crash frequencies on the before and after treatment that may be due to regression-to-the-mean. In accounting for regression-to-the-mean phenomenon, the number of crashes anticipated before installation of PCSs is a weighted mean of information from two sources. The first source is the number of crashes observed in the before period at intersections where PCSs have been installed. Additionally, the second source is the number of crashes predicted at signalized intersections with PCSs based on reference intersections without PCSs, which share similar traffic and physical characteristics.

To quantify the weights and the number of crashes anticipated on sites with similar traffic and physical characteristics, comparison intersections without PCSs but with similar traffic, and physical characteristics to the intersections with PCSs were used. This is similar in principle to the use of a comparison group in the comparison group method. Nonetheless, the two methods differ in one major aspect, i.e., for the before-and-after EB method, data from the reference intersections without PCSs are used to estimate safety performance functions (SPFs) that relate crash experience of the sites to their traffic and physical characteristics.

Safety performance functions

A safety performance function (SPF) is widely known as a crash prediction model that relates the crash frequency to traffic, geometric and other factors that influence the change in pattern and crash rates (Gross et al., 2010). It predicts the mean crash frequency for similar locations with the same characteristics mostly referred to as comparison sites. These characteristics typically include traffic volume, traffic control devices, geometric characteristics (number of lanes, road surface widths, shoulder widths, median characteristics), land use information and socio-demographic characteristics. Generally, comparison sites are used to account for time trends and changes in other factors such as traffic volumes and crash reporting systems. SPFs' development may employ two approaches, which are, simple SPF and full SPF. Full SPF is a mathematical relationship that relates all contradicting parameters that may influence variation in crash rate, including traffic and geometric parameters as predictor variables, while simple SPF includes AADT as the only independent variable in predicting crash frequency on a roadway facility (Gross et al., 2010). Thus, full SPFs are developed in this study, considering that they capture all contradicting attributes that influence the changes in crash frequency at the respective road entity, as crash frequency is not only affected by the traffic volume.

Choice of count model

Development of SPFs commences with the use of a count model to determine coefficients of model variables. In modeling crash counts, normally two categories of count modeling approaches are employed. These are Poisson and negative-binomial (NB) regression analysis. The choice between the two model types depends upon the relationship between the mean and the variance of the data in hand. Poisson regression analysis approach is employed in cases where the mean and variance of the data are equal. It is worth mentioning that, due to the possible positive correlation between observed crash frequencies, over-dispersion (variance of the data exceeds its mean) may occur (Hilbe, 2012). The HSM specifically calls for the use of the NB model in lieu of Poisson model. This is because the degree of over-dispersion in a negative binomial model is depicted by a statistical parameter normally called over-dispersion parameter. This parameter is estimated along with the coefficients of the regression equation (AASHTO, 2010).

In addition to the NB model, other models that account for overdispersion parameter include Poisson-lognormal, multivariate, hierarchical, Markov switching, Bayesian neural network, and support vector machine models (Lord & Mannering, 2010). It is also worth noting that the NB model is generally used because crash data have a mean which follow a gamma distribution and the variance of the crash data is normally greater than the mean (Shen, 2007). Hence, this study employs a NB regression analysis, shown in Equation 2.1, as described by Washington et al. (2003). The probability $P(y_i)$ of intersection i having N_i crashes in a given time period (yearly) is computed as follows.

$$P(y_i) = \frac{\Gamma((\frac{1}{\alpha}) + N_i)}{\Gamma(\frac{1}{\alpha}) N_i!} \left(\frac{1}{1+\alpha\lambda_i}\right)^{1/\alpha} \left(\frac{\alpha\lambda_i}{1+\alpha\lambda_i}\right)^{N_i} \quad (2.1)$$

Where

$\Gamma(x)$ a value of gamma function

N_i is the number of crashes for comparison intersections

λ_i designates the Poisson parameter for intersections without PCS (reference sites)

α is the over-dispersion parameter

$P(y_i)$ is the probability of intersection *i* having crashes N_i

Before-and-after with EB model formulation

The methodology established by Hauer was adopted to obtain EB estimates of the overall CMFs as well as Crash Modification Functions (CMFunctions) for different predictor variables (Hauer, 1997). The safety effectiveness of a treatment is estimated by comparing the observed number of crashes to the anticipated number of crashes after the installation of PCSs. The developed safety performance functions for each crash category was used to estimate the predicted crashes before the installation of PCSs, $\mathrm{N}_{\mathrm{predictedB}}$. Then the weighted adjustment factor (*W*) is computed using the total predicted crashes (by SPF) before the installation of PCSs and the over-dispersion parameter α for each crash category, as shown in Equation 2.2.

$$W = \frac{1}{1+\alpha \times \mathrm{N}_{\mathrm{predictedB}}} \quad (2.2)$$

Then, the number of expected average crash frequency before the installation of PCSs ($N_{expectedB}$) is computed using Equation 2.3:

$$\mathrm{N_{expectedB}} = W \times \mathrm{N_{predictedB}} + (1 - \mathrm{W}) \times \mathrm{N_{observedB}} \tag{2.3}$$

Where $\mathrm{N_{observedB}}$ refer to the total number of observed crashes at the treatment intersection

The expected average crash frequency in the after period for sites with PCSs assuming that PCSs were not installed is computed by multiplying the expected crashes before the installation of PCSs with *r* (Equation 2.4), where *r*, is the crash adjustment factor computed using predicted crashes (in SPF) after the installation of PCSs and before the installation of PCSs.

$$\mathrm{N_{expectedA}} = \mathrm{N_{expectedB}} \times r;\ \text{where} \qquad r = \frac{\mathrm{N_{predictedA}}}{\mathrm{N_{predictedB}}} \tag{2.4}$$

The CMF is then computed using Equation 2.5:

$$\mathrm{CMF}(\%) = \left(\frac{\frac{\sum \mathrm{N_{observedA}}}{\sum \mathrm{N_{expectedA}}}}{1 + \frac{\sum_{\text{all PCS intersections}} \mathrm{r}^2 \times \mathrm{N_{expectedB}} \times (1-W)}{\sum \mathrm{N_{expectedA}}^2}} \right) \times 100 \tag{2.5}$$

In addition, the CRF is computed using Equation 2.6:

$$\mathrm{CRF}(\%) = 1 - \mathrm{CMF} \tag{2.6}$$

Standard error of the CMF (ϑ) is then computed using Equation 2.7:

$$\vartheta = \sqrt[2]{\frac{\frac{1}{\sum \mathrm{N_{observedA}}} + \frac{\sum_{\text{all PCS intersections}} \mathrm{r}^2 \times \mathrm{N_{expectedB}} \times (1-\mathrm{W})}{\sum \mathrm{N_{expectedA}}^2}}{\left(1 + \frac{\sum_{\text{all PCS intersections}} \mathrm{r}^2 \times \mathrm{N_{expectedB}} \times (1-\mathrm{W}) \sum \mathrm{Var}}{\sum \mathrm{N_{expectedA}}^2}\right)^2} \times \mathrm{CMF}^2} \tag{2.7}$$

The 95% confidence interval is then computed using Equation 2.8. The CMF is considered significant when the 95% confidence interval does not include one (1.0) in its interval. This is because the CMF value of one (1.0) indicates no effect of the treatment on crash frequency at the treated site.

$$95\%\ \mathrm{CI} = \mathrm{CMF} \pm (1.96 \times \vartheta) \tag{2.8}$$

Model Results

The results discussed under this section are based on the CMFs for the safety impacts that PCSs have on drivers at signalized intersections. The CMFs were estimated using observational before-after study with empirical Bayes method. The SPFs used for predicting crash counts are discussed next.

Safety performance functions results

Florida-specific full SPFs for four-legged PCSs intersections on Major Street were developed. The SPFs were developed based on crash severity levels i.e., total (KABCO) crashes, fatal and injury (KABC), and property damage only (PDO). Additional SPFs were developed using different crash types; rear-end and angle crashes. Variables to be included into the SPF models were selected based upon their level of significance (*P*-values). In general, the SPF models had variables, which were significant at 95% and 90% level of confidence, with the exception of one (1) variable, which was significant at 85% level of significance. It is possible that the variable total number of lanes on the major approach is not that significant due to possible correlation with the traffic volume. The computed results for the five SPF models developed indicate the increase of crash frequency for intersections with higher traffic volumes. In addition, crash frequency was observed to be higher for intersections with higher speed limits (above 40 mph) and commercial land use areas.

CMFs for impact of PCSs to drivers

After computing the predicted crash counts from SPFs, the before-and-after study with empirical Bayes method was used to estimate the CMFs. The results of SPFs, CMFs, and CRFs are presented in Table 2.3. The results are provided for five crash categories, including crash severities (KABCO, KABC, and PDO), and crash types (rear-end and angle crashes). All crash categories have CMF less than one (1), indicating safety improvements on drivers' maneuverability at signalized intersections. Bolded CMF values for KABCO, PDO, and rear-end crashes are statistically different from 1.0 at a 95% level of confidence.

Table 2.3 SPFs, CMFs, and CRFs Results

	KABCO		**KABC**		**PDO**		**Rear-end**		**Angle**	
Parameters	Coef.	P>z	Coef.	P>z	Coef.	P>z	Coef.	P>z	Coef.	P>z
Over-dispersion	0.371		0.779		0.685		0.643		0.178	
Constant	-4.179	0.002***	-4.651	0.091**	-6.062	0.000***	-4.063	0.076**	-5.221	0.003***
Ln AADT (major)	0.156	0.045***	0.235	0.060**	0.331	0.029***	0.294	0.047***	0.331	0.030***
Ln AADT (minor)	0.147	0.033***	0.305	0.020***	0.440	0.009***	0.180	0.041***	0.244	0.082**
Total number of lanes (major)	0.459	0.100**	0.158	0.103*	0.129	0.111*	0.188	0.066**	0.195	0.102*
Total number of lanes (minor)	0.507	0.096**	0.508	0.020***	0.152	0.012***	0.240	0.040***	0.076	0.006***
Posted speed (major) > 40 mph	0.113	0.015***	0.117	0.096**			0.222	0.001***		
Commercial land use	0.105	0.013***								
CMF	**0.912**		0.952		**0.929**		**0.920**		0.954	
CRF (%)	8.8		4.8		7.1		8.0		4.6	
SE of CMF	0.029		0.079		0.034		0.016		0.080	

Note: Coef. = Coefficient, P>z= Level of significance whereby *** indicates significance at 95%, ** 90%, * 85%, SE=Standard Error, CMF= Crash Modification Factor whereby bolded CMF are significant at 95% level of confidence, CRF=Crash Reduction Factor.

The CMFs for different crash categories represent the expected changes in crashes for four-legged intersections with PCSs compared with the expected crashes on four-legged intersections without PCSs. The CMF for total crashes is 0.912, indicating a reduction in total crashes by 8.8%. This finding is consistent with the CMF obtained in a Michigan study (CMF of 0.946), estimated using the before-and-after with comparison group method (Kwigizile et al., 2015). The CMF for fatal and injury crashes is 0.952, with a percentage reduction in fatal and injury crashes due to the presence of PCSs by 4.8%. For the case of fatal and injury crashes, the Michigan study obtained a CMF of 0.927 (Kwigizile et al., 2015).

The difference in the CMFs between the ones obtained in this study and the ones in Michigan might be attributed to a number of factors. These include use of different CMFs estimation methods, whereby this study used before-and-after with EB method while the Michigan study used the before-and-after with comparison group method. Other factors may include differences in traffic, geometric, weather, and land use characteristics. The CMF for fatal and injury crashes was not significant at 95% level of confidence, the fact that may be attributed to a higher value of standard error for this crash category. For property damage only crash category, the obtained value of CMF is 0.929 indicates improvement in safety due to presence of PCSs (7.1% CRF).

Rear-end crashes were observed to be reduced by 8.0%, an improvement that is made due to presence of PCSs. Rear-end crashes are associated with unsafe stopping or reduction in speed of the leading vehicle. Information that PCSs provide to drivers enhances safety at signalized intersections since it enables drivers to have a prior decision when approaching signalized intersections. PCS timers provide drivers with an important cue, i.e., time remaining for their right-of-way at the intersection. This may, in turn, reduce the number of vehicles exposed to an intersection-approach dilemma zone. Consequently, drivers can start decelerating early if they realize that they cannot make the green, when PCS times approach zero, promoting comfortable and safe deceleration maneuvers when they are required to stop at the intersection.

Lastly, the CMF for angle crashes is found to be 0.954, indicating safety improvements of angle crashes by 4.6%. In addition, the CMF for this crash category was not significant at 95% level of confidence, due to a large standard error. Among other factors, angle crashes are influenced by illegal continuation to cross the intersection during the onset of red phase.

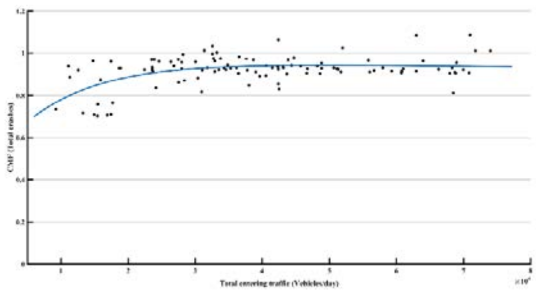
Estimation of CMFunctions

Crash modification functions explain the changes in safety effects while accounting for variation in geometric characteristics and other influential factors at the treated sites. Traffic volume (AADT) has been observed to be one of the prominent factors influencing the increase of crash frequency (Park et al. 2015). Thus, in this study, AADT was used as a continuous variable in developing the CMFunctions. Eventually, the relationship between the CMFs and AADT for each of the treated sites was developed. The developed CMFunctions are shown in Table 2.4.

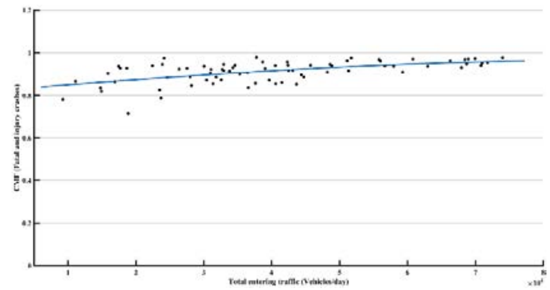
Table 2.4 CMFunctions for Different Crash Categories

Crash category	**Equation**	**R-square**	**Adjusted R-square**
KABCO	$\mathrm{CMF_i} = 0.969 \times \exp(\text{Total entering traffic } \times -0.000000462) + -0.4638 \times \exp(\text{Total entering traffic} \times -0.0000914)$	0.247	0.226
KABC	$\mathrm{CMF_i} = 12.88 \times \exp(\text{Total entering traffic } \times -0.00000772) + -12.05 \times \exp(\text{Total entering traffic} \times -0.00000803)$	0.333	0.305
PDO	$\mathrm{CMF_i} = 0.549 \times \text{Total entering traffic}^{0.051}$	0.323	0.313
Rear-end	$\mathrm{CMF_i} = -0.207 \times \exp(\text{Total entering traffic } \times -0.0000599) + 0.9475 \times \exp(\text{Total entering traffic} \times -0.0000000199)$	0.313	0.279

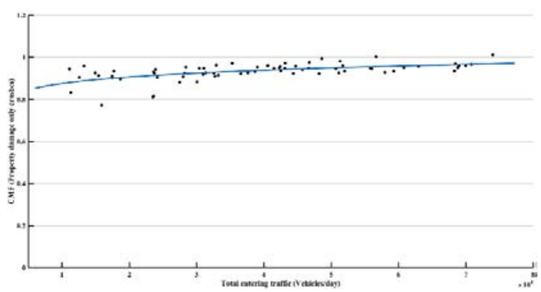
Non-linear regression models, i.e. inverse, power, quadratic, and exponential models, have been observed to be the best-fitted functions for different roadway characteristics (Park et al., 2015). As shown in Figure 2.2, linear and the abovementioned non- inverse, quadratic, power and exponential functions were compared, and the best-fitted functions for each crash category were selected based on the R-squared value. It is worth noting that, power and exponential functions were the best-fitted functions for this case. Treated sites were grouped based on the frequency distribution of crashes for each crash category in a manner to avoid observations with zero crash counts.



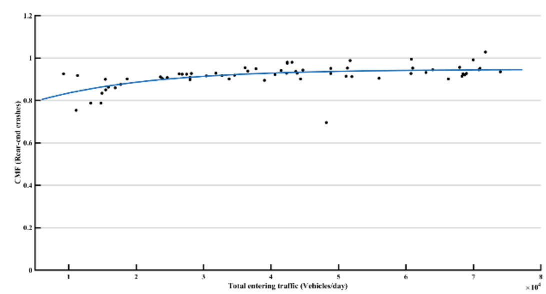
2(a) Total crashes CMFunctions



2(b) Fatal and injury crashes CMFunctions



2(c) PDO crashes CMFunctions



2(d) Rear-end crashes CMFunctions

Figure 2.2. Crash modification function curves

Conclusions and Recommendations

In view of the fact that pedestrian countdown signals (PCSs) are meant to help pedestrians cross the intersection safely, the same signals could give cues to drivers as they approach the intersection. This study focused on evaluating the safety effectiveness of PCSs to drivers at signalized intersections using the state maintained intersections in Florida (cities of Gainesville and Jacksonville). In total, 110 sites with PCSs and their respective comparison sites (without PCSs), 93 in total, were examined, using data collected from years 2003 through 2014.

The before-and-after empirical Bayes with comparison group method was employed in developing the CMFs considering its ability to produce reliable and robust estimates. The CMFs were developed based on crash severity including total, fatal/injury (F+I), and property damage only (PDO) crashes. In addition, the study developed CMFs for two crash types, i.e. rear-end and angle crashes. Moreover, to observe the relationship between the CMFs and traffic volume, crash modification functions (CMFunctions) for different crash categories used in this study were developed.

Full safety performance functions for each of the aforementioned crash categories were developed from comparison intersections based on heterogeneous factors. These include additional factors that influence changes in crash frequency and severity patterns at the treatment sites independent of the installed treatment, PCSs for this case. The heterogeneous factors incorporated in this study include traffic volume, geometric characteristics, traffic conditions, and adjacent land use.

According to the CMFs obtained in this study, PCSs have safety benefits to drivers. The results suggest that cues provided by PCS' timers to drivers as they approach intersections may help them in reducing conflicts that may lead to rear-end and angle crashes. In summary, this study demonstrated that apart from aiding pedestrians, installation of PCSs improves the safety performance of drivers at signalized intersections.

Moreover, the CMFunctions developed from this study prove that CMFs vary at different treatment sites and with distinct roadway characteristics, area type, socio-economic characteristics, and time. Hence, more work is required to improve the developed CMFunctions by incorporating more roadway and other pertinent characteristics. It is worth mentioning that the developed CMFs can be applied in other areas apart from the study sites by applying respective location calibration factors.

Unavailability of traffic volume data especially on non-state maintained roadways limited the site selected for the study to state-maintained intersections only. Further, because the minor street roadways for some of the intersection are mostly non-statewide maintained roadways, they also lack reliable traffic volume. This in turn, resulted in dropping of many sites from the analysis due to incompleteness of data. In addition, the process of retrieving information on the installation dates of PCSs was long and tedious. This was due to the absence of a database with the dates of PCSs installation. Maintaining a database with the records of the installation dates for traffic control devices such as PCSs is necessary to aid continued efforts in evaluating the effectiveness of such devices. To the authors' knowledge the only geometric change that was made on the treatment intersections is the installation of PCSs. It was not possible to collect information on the signal retiming changes made on the study intersections during the study period due to frequent retiming efforts that agencies undertake once they realize a change in traffic patterns.

CHAPTER 3: PAPER 2

PAPER II: A full Bayesian Approach to Appraise the Safety Effects of Pedestrian Countdown Signals to Drivers

Paper II is prepared to be submitted to the journal of Accident Analysis and Prevention.

Introduction

Each year, more than 32,000 fatalities and 2 million non-fatal injuries occur on United States roadways. These tragedies amount to an estimated societal burden of more than $230 billion of medical and other costs (Blincoe et al., 2015; NHTSA, 2010). The U.S has experienced a 31 percent decrease in its motor vehicle fatality rate per capita over the past 13 years. Even so, compared with 19 other developed countries, which experienced on average a 56 percent reduction in the frequency of fatal crashes during the same period, the U.S has the slowest reduction (31%) (Sauber-Schatz et al., 2016). Shockingly, latest data from the National Highway Traffic Safety Administration (NHTSA) indicate a 7.2 percent increase in roadway fatalities in 2015, shooting from 32,744 in 2014 to 35,092 in 2015 (NHTSA, 2016). This amounts to nearly 700 deaths every week due to traffic collisions. To put these statistics in perspective, a number of lives lost due to roadway crashes in the U.S. is equivalent to two commercial large aircrafts, such as the Airbus A340 500 (capacity of 372 seats), crashing every week.

Although they include a small proportion of the overall roadway network, compared to other roadway segments, intersections are characterized by increased conflicts due to various conflicting traffic movements converging at the same location. The U.S. Department of Transportation estimates that 43 percent (43%) of motor-vehicle crashes occur at intersections or are intersection-related. In some cases, the conflicts at intersections involve more than one transportation mode as drivers, pedestrians and cyclists come across at the same point.

Pedestrians are considered vulnerable road users, hence ensuring their safety when crossing intersections is of paramount importance. The Pedestrian countdown signals (PCSs) are conventionally installed to improve pedestrian safety at signalized intersections. Generally PCSs, through the timer, are used to show the remaining seconds for pedestrians to cross the intersection during the pedestrian clearance interval. There is ample research evidence that shows safety benefits of PCSs for pedestrians (Huang and Zeeger, 2000; Markowitz et al., 2006; Chen et al., 2015; Lambrianidou et al., 2013; Schmitz, 2011; Scott et al., 2012; Vasudevan et al., 2011; and Eccles et al., 2004).

Despite being intended for pedestrians, the same information offered by PCSs to pedestrians has been observed to give cues to drivers as well. A few studies have documented on the effect of PCSs to drivers. These studies have mostly concentrated on the operational and capacity effects of these signals, such as the studies by Nambisan & Karkee (2010), Schmitz (2011), and Elekwachi (2010). The literature on the safety effectiveness of PCSs on drivers is scarce.

A literature search uncovered only two studies, both recent, that evaluated the safety effectiveness of PCSs on drivers. The first study (Kwigizile et al., 2015) was conducted in Michigan using the before-and-after with the comparison group method. According to the study, the presence of PCSs at signalized intersections reduce five percent (5%) of total crashes for all drivers. This finding was in line with a Florida study (Kitali et al., 2017) that employed the same method and observed 8.2 percent (8.2%) reduction in total crashes. This indicates that drivers utilize information provided by PCS timers to make informed decisions when approaching and crossing signalized intersections. Both studies employed the empirical Bayes before-and-after technique, which suffers from methodological and statistical limitations, including small sample size, inability to account for the uncertainty of the computed regression coefficients from the safety performance function (SPF) into the odds ratio computations (two-step procedure), and inability to develop multivariate models. These limitations can be potentially addressed by employing a Full Bayes (FB) method in lieu of a conventional empirical Bayes. While the before-and-after EB method has been the most acceptable technique for evaluating safety effectiveness of various roadway countermeasures since the inception of the first Highway Safety Manual (HCM, 2009), there has been an increased use of Full Bayes (FB) before-and-after technique for safety study over the last few years.

A FB approach has the ability to account for most of the uncertainties in the dataset and model parameters and thus overcome the maximum likelihood methods' problem of overestimating precision because of ignoring this uncertainty (Park et al., 2016). The FB methodology is also a single-step integrated procedure, i.e. it integrates the process of estimating the SPF and treatment effect in a single step, thus incorporates the uncertainties of the SPFs in the final estimates. The FB methodology is independent of sample size, thus yielding robust results even when used with small sample size (Li et al., 2013; Ahmed, et al., 2015). Another important advantage of FB approach is the ability to allow inference at more than one level for multilevel (multivariate) models. The multivariate approach takes into account the fact that crash data of different severities e.g. property damage only (PDO) and fatalities and injury (FI) crashes are correlated. Unlike the negative binomial model that is widely used in the EB methodology, the FB approach makes use of hierarchical models i.e. Poisson-Gamma and Poisson-lognormal distributions (Miaou & Lord, 2003; Lan et al., 2009; Pawlovich et al., 2006). Additionally, the FB approach divides the periods into time intervals (yearly in this case) and models, each time interval as a separate data point to account for time variations, unlike the EB methodology which average the data into a single data point.

Considering the novelty of the FB methodology and the fact that none of the previous studies has quantified the effects of PCSs to drivers using this approach, a study is therefore warranted.

Research Objective

The main purpose of this study is to evaluate the safety effectiveness of PCSs to drivers. The study employs the FB methodology to evaluate the overall safety effectiveness in terms of effectiveness for specific types of crashes – rear-end and angle crashes, in particular, and injury severity. The study also reported at the time-based effectiveness, due to the flexibility of the FB approach in analyzing changes of treatment effectiveness with time, for the after- period.

Study Significance

The findings of my thesis can be used by transportation agencies as part of their decision making when deciding on installing PCSs at signalized intersections. Transportation officials can incorporate the crash modification factors developed in this study to conduct an economic appraisal of installing PCSs. Also, the findings of this study may prompt a need for a much broader research to investigate whether the design of PCSs should target both pedestrians and drivers. This could lead to a fundamental change in the design of PCSs to enable not only pedestrians but the drivers as well to see the information displayed on PCSs at a reasonable sight distance– e.g., size can be increased to allow drivers approaching from a distance to observe the information on the timer in advance. Further, the height, angle and location of the countdown timer can be adjusted to allow drivers in multiple lanes to see them in advance. Also, the research community could use the this study's methodology as a building block to support a broad, ongoing effort aimed at mainstreaming the use of FB approach for evaluating safety improvement projects.

Background

Full Bayes methodology

Historically, crash prediction models have employed mainly maximum likelihood models (Hauer, 2001; Park et al., 2015; Persaud & Lyon, 2007). Even after the introduction of the first version of the Highway Safety Manual (AASHTO, 2009), which advocated the use of empirical Bayes, model coefficients were still determined based on maximum likelihood models, the negative binomial being a preferred one. Recently, there has been a substantial increase in the use of hierarchical Bayesian approach in crash modeling. This increase can be attributed to a number of aspects including the availability of open source scripting software packages and the invention of strong computers that can perform complex statistical iterations such as Markov Chain Monte Carlo (MCMC) simulations. The use of FB in crash predictions dates more than two decades ago (Schlüter et al., 1997). But is was only at the end of the last decade that highway safety modeling scholars have increasingly researched the use of the FB approach applying the MCMC simulation (Aul & Davis, 2006; Carriquiry & Pawlovich, 2004; Davis & Yang, 2001; El-Basyouny & Sayed, 2009b; Lan et al., 2009; Li et al., 2008; Miaou & Lord, 2003; Park & Lord, 2007; Park et al., 2010; Pawlovich, et al., 2006; Persaud et al., 2010; Sacchi et al., 2015). It is worth mentioning that it was only about a decade that the hierarchical Poisson regression models with a change point to before-and-after evaluation were introduced in the arena of the FB technique (Aul & Davis, 2006; Davis & Yang, 2001; El-Basyouny & Sayed, 2009b; Lan et al., 2009; Li et al., 2008; Park & Lord, 2007; Park et al., 2010; Pawlovich, et al., 2006; Persaud et al., 2010; Sacchi et al., 2015).

Unlike the classical statistical theory, Bayesian statistics use the density function to estimate the effect of a given parameter on the model rather than a discrete coefficient (Ntzoufras, 2009; Saito et al., 2011). Use of the density function permits for a better understanding of the amount of uncertainty in the data, where the density function for each parameter provides the likelihood pertaining to a certain prediction effect (Saito et al., 2011). In Bayesian statistics, all unknown parameters are considered as random, thus requiring the definition of prior distribution initially.

The Bayesian technique incorporates prior information and observed information to develop an estimate for the expected crashes of the sites of interest, intersections with PCSs for this case. In the context of the crash prediction modeling, the prior information is the anticipated crash frequency from comparison locations and the observed information are the observed crashes on the treatment sites before the installation of the treatment (Persaud et al., 2010).

After observing the total number of crashes occurring in the particular study intersections, posterior distribution $\pi(\theta \mid y)$, described in Equation 3.1, is estimated. This computation is done by combining the *priori* and the observed data. It is worth mentioning that this posterior distribution is the key element in Bayesian inference (Ntzoufras, 2009; Saito et al., 2011). To account for uncertainty associated with crash modeling, the probability of θ, given y or($f(\theta \mid y)$) needs to be computed. This probability is used to make inference about the entire population.

$$\pi(\theta \mid y) = \frac{f(y|\theta)\pi(\theta)}{P(y)} \tag{3.1}$$

Where:

y	=	Crash frequency, and
θ	=	Vector for the unknown modeling parameters (priori)
$\pi(\theta \mid y)$	=	posterior distribution of θ, determined from known crash data,
$f(y\|\theta)$	=	likelihood of y given θ, and
$\pi(\theta)$	=	informational prior distribution of θ

P(y) is estimated through integrating $f(y|\theta)\pi(\theta)d(\theta)$, thus making the Bayesian equation a complete density function (Saito et al., 2011). Hence the final equation will be as described in Equation 3.2:

$$\pi(\theta \mid y) = \frac{f(y|\theta)\pi(\theta)}{\int f(y|\theta)\pi(\theta)d(\theta)} \tag{3.2}$$

The denominator in Equation 3.2 needs to be integrated for each of the parameters since the model composes more than one parameter (Saito et al., 2011). Due to the increasing complexity of this equation, a statistical technique such as the Markov Chain Monte Carlo (MCMC) can be useful in computing such a complex integration. With an appropriate number of samples that allow the model to converge, the true posterior distribution can be accurately estimated.

Markov Chain Monte Carlo simulation in full Bayes studies

MCMC simulation is a stochastic simulation technique that is useful for computing inferential quantities. It has been widely employed in Bayesian statistics due to the complexity of the integration needed for approximating posterior distributions. There are different types of MCMC algorithms. The three most widely used are Gibbs sampler (WinBUGS software) (Sacchi et al., 2015), Metropolis-Hastings algorithm (MATLAB) (Park et al., 2010), and Hamiltonian Monte Carlo (HMC) algorithm (R software). The fact that this algorithm employs the use of physical system dynamics rather than a probability distribution to estimate future states in the Markov chain make it appealing over the other two MCMC algorithms (Brooks et al., 2011). This is because the use of physical system dynamics allows the Markov chain to approach the target distribution more efficiently and thus resulting in faster convergence.

FB approach improvements on the SPF development

The abovementioned benefits of the FB approach over other safety effectiveness methodologies including EB allows additional flexibility in the development of the crash prediction model (SPF). In the FB methodology, prior information and observed data are combined to develop a single robust statistical model which is used to generate a posterior distribution on which inference on selected parameters can be based. The hyper-prior distributions defined while estimating the posterior distribution for the anticipated number of crashes is carried over throughout the modeling process and finally the safety effectiveness computations. Conversely, the EB approach employs the use of an external function, SPFs, to derive the parameters of prior distributions for the predicted crashes, and consider them as true parameters once they are estimated. Ostensibly, the associated uncertainties in the regression model parameters of SPFs are not included in the final safety effectiveness estimate (Park et al., 2010).

The FB approach has the capability to accounting for different variations and characteristics existing in the crash data such as the use of intervention models during evaluation of the safety effects of the installed countermeasure on a road (El-Basyouny & Sayed, 2011; Chen & Persaud, 2014; Li et al., 2008; Park et al., 2010; Pawlovich et al., 2006). An intervention model allows for the exploration of trends that may occur in between the before- or after- periods. This model also allows for the investigation of the temporal effects of traffic safety under the hypothesis that its effect changes over time as opposed to occurring instantaneously.

Another important flexibility brought about by the flexibility of the FB approach is the application of the multivariate Poisson lognormal (MVPLN) to model number of crashes at different severity levels (El-Basyouny & Sayed, 2009b; El-Basyouny & Sayed, 2011). The FB approach also allows the incorporation of random parameters to account for the unobserved heterogeneity. (El-Basyouny & Sayed, 2009a; Li et al., 2008).

Jump parameter

Crash frequency for treatment sites is subject to change due to the effect of the installed treatment. Given that changes may not be gradual, an immediate drop or increase in crash frequency is expected at the respective sites after the intervention. The model parameter that accounts for the immediate drop or increase in the crash frequency at the treatment sites is conventionally referred to as a jump parameter. It has been incorporated in several studies including (Li et al. 2008; Li, et al., 2013).

Random parameters

As the name suggests, the probability distribution of the Poisson mean for the Poisson lognormal count model is lognormal. This probability is associated with the independent variables *x* by regression coefficients. Independent variables included in the regression models are variables observed by the analyst from the historical data. Considering that crashes are rare and random events, the observed independent variables cannot address all of the heterogeneity existing in the crash occurrence events. Random parameters are included in the model to account for heterogeneity, which are unobserved factors that may vary across observations. Studies that have introduced random parameters in count models have reported improved prediction accuracy (Anastasopoulos & Mannering, 2009; Li et al., 2008).

Cross-validation as a measure of models' prediction accuracy

Cross-validation (CV) is normally employed to evaluate the future predictive capability of the models developed under different simulation environments including Bayesian (Xie et al., 2014; Yang et al., 2013). There are different types of cross-validation methods are used to assess the predictive performance of Bayesian models, the Holdout CV method being one of the simplest (Arlot & Celisse, 2010). In holdout CV, dataset used in the analysis is randomly divided into two sets, namely the training and the testing set. The training set is used in model fitting, and then the testing dataset is used to beta test the performance of the model. The shortcoming of this technique is that the output of this method depends on the distribution of the data points on the two sets i.e. training set and testing dataset. Further, this technique is limited to larger datasets only.

One way to overcome the shortcomings of the holdout CV method is to use the K-fold CV (Kuhn & Johnson, 2016). In this method, the dataset is separated into *k* subsets, and the holdout method is repeated *k* times (Akay et al., 2015). The variance of the resulting estimate decrease as k is increased. The main setback of this method is that selection of number of subsets (k) is biased.

Leave-one-out (LOO) cross validation technique has proven to be most reliable over other CV including K-fold CV (Stan Development Team, 2016). This approach computes out-of-sample prediction accuracy from a fitted Bayesian model using the log-likelihood estimated at the posterior simulations of the parameter values (Vehtari et al., 2016). LOO has numerous benefits over the simpler estimates of predictive error such as Alkaike Information Criterion (AIC) and Deviance Information Criterion (DIC) but are less used in practice due to intense computational requirements.

Data Collection

The dataset considered in this study was extracted from different four-legged signalized intersections in Jacksonville and Gainesville, Florida. The installation date of PCS on the treatment sites ranges between the years of 2006 through 2011. For each of the treatment sites, three (3) years before the installation of PCSs and three (3) years after installation of PCSs were used for analysis of changes in crash frequency due to the installation of PCSs. It is worth mentioning that the respective year that PCS was installed in each of the sites was excluded from the study to allow enough buffer time for changes brought about by PCSs. One hundred and ten (110) treatment intersections and 93 comparison intersections were selected. Comparison intersections were selected according to their geographical proximity and similarity to the treatment sites in terms of traffic and geometric characteristics.

Data were collected and retrieved from the following databases; Florida Unified Base-Map Repository (crashes), Florida Geographic Data Library (FGDL) Metadata explorer (land use information), and FDOT GIS database (posted speed). Other sources of data included Google earth-street view and historical imagery tool to retrieve geometric information from previous years before and after installation of PCSs and Florida Traffic Monitoring Sites (TMS) (traffic volume for the year 2003 through 2014). The historical imagery tool in Google earth Pro software was utilized in ensuring the quality of the developed SPF by checking the reference sites to verify that there is no major geometric change during the study period. Table 3.1 summarizes the crash types used in this study in terms of their means and standard deviation (SD).

Table 3.1 Annual Crash Data Summary: Treatment and Comparison Intersections

No.	Type of crash	Intersection type	Before		After	
			Mean	SD	Mean	SD
1	Total	Treatment	18.38	17.94	14.31	14.56
		Comparison	19.97	20.4	13.23	12.25
2	F + I	Treatment	11.73	11.79	10.55	9.56
		Comparison	10.79	9.27	11.03	10.22
3	PDO	Treatment	10.5	12.93	9.3	10.08
		Comparison	10.91	10.34	9.06	11.93
4	Rear-end	Treatment	10.09	12.76	6.74	8.07
		Comparison	11.65	10.44	7.23	8.57
5	Angle	Treatment	5.29	5.47	4.75	5.16
		Comparison	5.00	5.09	3.78	4.82

Note: SD= standard deviation.

Important variables considered for this study includes traffic volumes on the major and minor approaches, geometric characteristics of major and minor approaches, and the area type (commercial categorized as 1 and other land use types as 0) as illustrated in Table 3.2. Other information retrieved on the major street include posted speed (>40 mph coded as 1, otherwise 0).

Table 3.2 Selected Variables in Treatment Sites

	Variable	Mean	Std. Dev.
	Land use (1 commercial, otherwise 0)	0.47	0.50
Major	Average AADT (Vehicle/day)	28894.20	12281.44
	Total number of lanes	5.33	0.95
	Posted speed (mph)	41.18	6.74
Minor	Average AADT (Vehicle/day)	11427.93	8120.44
	Total number of lanes	2.93	1.00

Note: AADT=annual average daily traffic.

Poisson-Lognormal model: A Statistical Model to Quantify the Impact of PCSs to Drivers

In this study, the Poisson lognormal statistical model is considered to assess the effect of the intervention. The Poisson-lognormal model derivation in this study is derived from an extensive literature (Li et al. 2008; Zhou, et al., 2012). In all cases, Y_{it} denotes the crash count observed at site $i(i = 1,2,3, \ldots, n)$ during year t $(t = 1,2,3, \ldots, m)$, and

$$Y_{it}|\theta_{it} \sim Poisson\ (\theta_{it}), \tag{3.3}$$

The Poisson mean θ_{it} can be written as shown in Equation 3.4

$$\theta_{it} = \mu_{it} = e^{(\beta_0 + \ldots + \beta_n X_n)} \tag{3.4}$$

It is worth noting that, the Poisson model assumes the mean and the variance of the crash counts are equal, Equation 3.5.

$$E[\mathrm{Y_{it}}|\mathrm{\theta_{it}}] = Var[\mathrm{Y_{it}}|\mathrm{\theta_{it}}] = \mathrm{\mu_{it}} \tag{3.5}$$

In practice, however, crash data are often over-dispersed, considering the heterogeneity partly contributed by the rareness and randomness of these events. To incorporate the over-dispersion attribute of the crash data, the Poisson regression model can be modified as in Equation 3.6, where ε_i is a multiplicative random-effect that model heterogeneity across individual crashes.

$$\theta_{it} = \mu_{it} + \varepsilon_{i} = e^{(\beta_0 + \ldots + \beta_n X_n)\varepsilon_i} \tag{3.6}$$

Also, the mean and variance for the new Equation 3.6 will be as expressed in Equations 3.7 and 3.8 respectively, where the variance is greater or equal to the mean.

$$E[Y_{it}|\theta_{it}] = e^{(\beta_0 + \ldots + \beta_n X_n)\varepsilon_i} + E[\varepsilon_i] \tag{3.7}$$

$$Var[Y_{it}|\theta_{it}] = E[Y_{it}|\theta_{it}] + \frac{Var[\varepsilon_i]}{E^2[\varepsilon_i]} E^2[Y_{it}|\theta_{it}] \tag{3.8}$$

Using the Poisson lognormal, the random effect ε_i will be lognormally distributed as presented in Equation 3.9.

$$\varepsilon_i \sim lognormal(0, \sigma^2) \tag{3.9}$$

Based on expectation of the mean, $E[\varepsilon_i] = e^{\sigma^2/2}$ and the variance, $Var[\varepsilon_i] = e^{\sigma^2}\left(ln^{\sigma^2} - 1\right)$ using Equation 3.7 and 3.8, the modified Equations are expressed in Equations 3.10 and 3.11, respectively.

$$E[Y_{it}|\theta_{it}] = e^{(\beta_0 + \ldots + \beta_n X_n + \varepsilon_i)} \tag{3.10}$$

$$Var[Y_{it}|\theta_{it}] = E[Y_{it}|\theta_{it}] + \left(e^{\sigma^2} - 1\right)E^2[Y_{it}|\theta_{it}] \tag{3.11}$$

Then the marginal mean and variance of Y_{it} will be as described in Equations 3.12 and 3.13 respectively.

$$E(Y_{it}) = E[E(Y_{it}|\theta_{it})] = \mu_{it}, \tag{3.12}$$

$$Var(Y_{it}) = E[V(Y_{it}|\theta_{it})] + V[E(Y_{it}|\theta_{it})] = \mu_{it} + e^{(\sigma^2)-1}{\mu_{it}}^2 \tag{3.13}$$

Estimation of the Posterior Distributions of Count Model Parameters

In this study, crash modification factors are used to assess the effectiveness of PCSs to drivers. The study uses the before- and after- method, employing crash data collected before and after installing PCSs. Let Y_{it} denote the crash count observed at site i $(i = 1, 2, 3, \dots, n)$ during year t $(t = 1, 2, 3, \dots, 6)$. To incorporate the linear intervention model, let T_i signify the treatment indicator (equals 1 for treatment intersections, 0 for comparison intersections), t_{0i} represent the interventions year for the i[th] treatment intersections and its matching comparison intersections, I_{it} indicate the time indicator (equals 1 in the after period, 0 in the before period). For exposure variables let V_{1it} and V_{2it} denote the annual average daily traffic (AADT) on the major and minor approaches respectively.

In addition, let $(X_{3i}, \dots, X_{ji})$ symbolize other explanatory variables including geometric and land use characteristics. These includes number of lanes on the major and minor approaches, land use information, and speed limit.

Model 1: Poisson-lognormal model with individual site random effect

The lognormal model for crash density is described as a piecewise linear function (Equation 3.14) of predictor variables, such that the function is continuous at the change point t_{0i}. The piecewise linear function is defined by at least two equations, each of which applies to a different part of the domain i.e. before and after installation of the PCSs in this case. The site-level random effect α_i is also included as shown in Equation 3.14.

$$ln(\mu_{it}) = \alpha_0 + \alpha_1 T_i + \alpha_2 t + \alpha_3 (t - t_{0i}) I_{t > t_{0i}} + \beta_1 V_{1it} + \beta_2 V_{2it} + \dots + \beta_n X_n + \alpha_i \quad (3.14)$$

Where $\alpha_i \sim N(0, \sigma_\alpha^2)$ and σ_α^2 accounts for the variation existing between intersections in yearly log crash frequency. The linear-intervention model allow for different slopes of crash frequency on time before and after the installation of the PCSs and also across the treatment and comparison intersections

Model (1) in Equation (3.11) is limited to account for potential difference across the treatment and the comparison intersections. Considering the crash data were collected for six years, i.e. three years before the installation of the treatment and three years after the installation of the treatment, it is also possible that there is a variation of the slope across the data for different years. Thus, a parameter, which will account for the heterogeneity of the crash data collected at different time periods is required to improve the reliability of the model estimates.

Model 2: Poisson-lognormal model with individual site random effect and jump parameter

Considering the fact that there might be an immediate drop or increase in crash frequency upon installation of an intervention, it is worth incorporating a parameter that will account for this scenario, the jump parameter for this case. Equation 3.15 is an improvement of Equation 3.14 with an addition of the jump parameter, α_4.

$$\ln(\mu_{it}) = \alpha_0 + \alpha_1 T_i + \alpha_2 t + \alpha_3 (t - t_{0i}) I_{t>t_{0i}} + \alpha_4 T_i I_{t>t_{0i}} + \beta_1 V_{1it} + \beta_2 V_{2it} + \cdots + \beta_n X_n + \alpha_i \tag{3.15}$$

This additional parameter, α_4, in the log link function of the Poisson regression model is also normally distributed and independent of other parameters.

Model 3: Poisson lognormal model with individual-site random effect, jump parameter, and pair-random effect

The design of the FB before-and-after with comparison sites study includes comparison sites pairs for each of the treatment site. The comparison intersections were selected in a way to entail comparable geometric, traffic, and land use characteristics to the treatment intersections, a possible correlation between the crash frequencies of the two groups of sites maybe induced. In order to account for a possible correlation across the treatment-comparison intersection pairs, a random effect parameter δ_k for the kth pair is introduced in Equation 3.15, creating a new equation (Equation 3.16).

$$\ln(\mu_{it}) = \alpha_0 + \alpha_1 T_i + \alpha_2 t + \alpha_3 (t - t_{0i}) I_{t>t_{0i}} + \alpha_4 T_i I_{t>t_{0i}} + \beta_1 V_{1it} + \beta_2 V_{2it} + \cdots + \beta_n X_n + \alpha_i + \delta_k \tag{3.16}$$

Where, $\delta_k \sim N(0, {\delta_k}^2)$ accounts for the variability between the treatment intersection and the comparison intersection within each pair

Figure 3.1 demonstrates, in summary, the procedure used to select the model with the highest prediction accuracy (LOO CV). In this study, the approximate LOO cross-validation approach was employed using the Pareto-smoothed importance sampling (PSIS) method, a new procedure for regularizing important weights. PSIS is an advanced method, which provides improved accurate and reliable CV estimate by fitting a Pareto distribution to the upper tail of the distribution of the important weights (Vehtari et al., 2016).

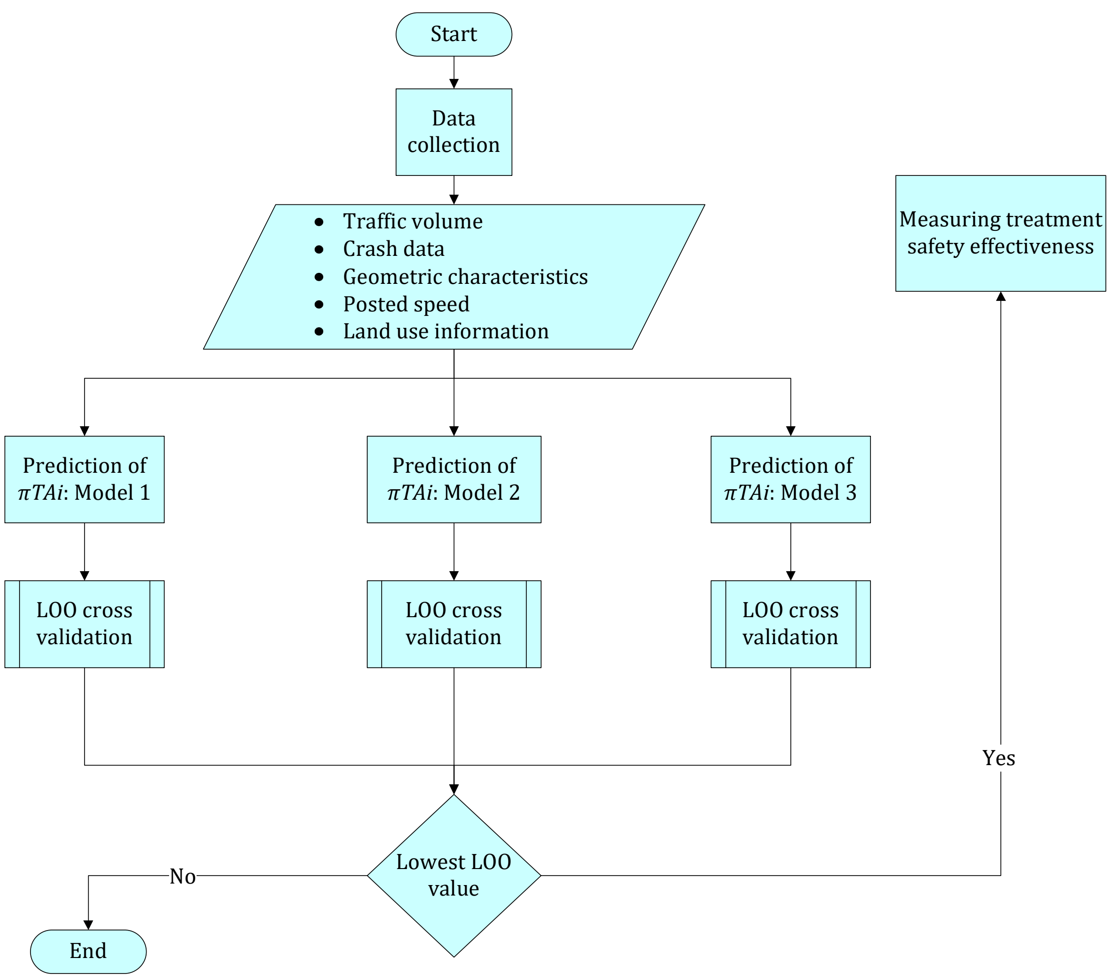


Figure 3.1. Full Bayes methodology flowchart

The Bayesian LOO estimate of out-of-sample predictive fit is:

$$\mathrm{elpd_{loo}} = \sum_{i=1}^{n} \log(p(y_i|y-i)) \tag{3.17}$$

Where

$\mathrm{elpd_{loo}}$ = LOO expected log predicted density

$p(y_i|y-i) = \int p(y_i|\theta)p(\theta|y-i)d\theta$ = leave-one-out predictive density given the data without the i^{th} data point

It is worth noting only the best model was used in quantifying the treatment safety effectiveness.

Measuring Treatment Effectiveness

Let μTBi and μTAi represent the predicted crash counts for the i^{th} treatment intersection averaged over three years of the before years and after years, respectively. Moreover, let μCBi and μCAi denote the corresponding counts for the paired comparison intersections. Let the superscripts T and C denote treatment and comparison intersections, respectively, and let the superscripts A and B denote the after and before periods, respectively. The ratio $\frac{\mu_{CAi}}{\mu_{CBi}}$, conventionally known as comparison ratio, is included during evaluation of the safety effect of the countermeasure to account for external confounding factors that might influence the change in the crash frequency apart from the installed treatment. These include improvement in the vehicle safety technology, new traffic policies, and traffic safety awareness education, among other factors. Explicitly, the estimate of this ratio is combined with the observed crashes on treatment intersections during the before period to compute the expected crashes on the treatment intersections had the PCSs not installed. In summary, the following steps were employed when implementing the FB before-and-after with comparison intersections method in evaluating the safety impacts of PCSs to drivers at signalized intersections.

Step 1:

Step 1 involved specification of the hyper-parameter values for prior distribution of model parameters, i.e. intercept and coefficient priors ~ normal (location (mean), scale (standard deviation)). The draw of model parameters and the expected annual crash frequency was obtained for each intersection *(i)*, and year *(t)* by MCMC, using the HMC algorithm for this case.

Step 2:

Next, the posterior distributions of crash frequencies during the before period for the comparison intersections μ_{CB} were estimated in line with the crash frequencies for the comparison intersections during the after period μ_{CA}, using Equations 3.18 and 3.19 respectively.

$$ln({\mu_{it}}^{CB}) = \alpha_0 + \alpha_2 t^B + \beta_1 V_{1it} + \beta_2 V_{2it} + \cdots + \beta_n X_n \qquad (3.18)$$

$$ln({\mu_{it}}^{CA}) = \alpha_0 + \alpha_2 t^A + \alpha_3 (t^A - t_{0i}) I_{it} + \beta_1 V_{1it} + \beta_2 V_{2it} + \cdots + \beta_n X_n \qquad (3.19)$$

Step 3:

To obtain the comparison ratio, the logarithm of the expected crash frequency for before period is subtracted from the logarithm of the expected crash frequency during the after periods shown in Equation 3.20.

$$R_{Ci} = ln({\mu_{it}}^{CA}) - ln({\mu_{it}}^{CB}) \qquad (3.20)$$

Step 4:

Subsequent to estimation of the comparison ratio, the predicted crash frequencies in the after period for each of the treatment sites had the PCSs not installed $\ln(\pi_{it}^{TA})$ can be computed. Using the comparison ratio R_{Ci} and the expected crash frequency for the treatment site during the before period computed using Equations 3.21, $\ln(\pi_{it}^{TA})$ is estimated as shown in Equation 3.22.

$$ln(\mu_{it}{}^{TB}) = \alpha_0 + \alpha_1 T_i + \alpha_2 t^B + \alpha_4 T_i t^B + \beta_1 V_{1it} + \beta_2 V_{2it} + \cdots + \beta_n X_n \tag{3.21}$$

$$ln(\pi_{it}^{TA}) = ln(\mu_{it}{}^{TB}) + R_{Ci} \tag{3.22}$$

Step 5:

Finally to obtain the safety effectiveness of the PCS to drivers, the logarithm of the expected crash frequency on the treatment sites during the after periods were computed (Equation 3.23). The PCS CMF, θ_{it}, was then computed based on the definition of the odds ratio i.e. the ratio of the observed crashes on the treatment intersections following the installation of PCS to the predicted crashes on the treatment sites assuming PCS were not installed during the after period (Equation 3.24).

$$ln(\mu_{it}{}^{TA}) = \alpha_0 + \alpha_1 T_i + \alpha_2 t^A + \alpha_3 (t^A - t_{0i}) I_{it} + \alpha_4 T_i t^A + \alpha_5 T_i (t^A - t_{0i}) I_{it} + \beta_1 V_{1it} + \beta_2 V_{2it} + \cdots + \beta_n X_n \tag{3.23}$$

$$\theta_{it} = ln(\mu_{it}{}^{TA}) - ln(\pi_{it}^{TA}) \tag{3.24}$$

Subsequent to θ_{it}, its posterior distribution was also computed following the Bayesian principles. The point estimates for CMF with their respective uncertainty estimates (standard deviation) were retrieved as the sample means of the respective posterior distributions of θ_{it}. Ultimately, the 95% Bayesian credible intervals (BCI) of CMF was also constructed from the estimated posterior distributions using the 2.5th percentile and the 97.5th percentiles.

Model Estimation using Hamiltonian Markov Chain (HMC) Algorithm

The Bayesian analysis performed in this study employs the Hamiltonian Monte Carlo (HMC) algorithm. The HMC is an MCMC algorithm that uses the derivatives of the density function being sampled, Poisson lognormal for this case, to produce the posterior distributions of the parameters intended to be estimated. It employs the principles of the Hamiltonian dynamics simulation that is based on numerical integration. For the case of MCMC simulation, the Hamiltonian dynamics simulation is corrected by performing the Metropolis acceptance step (Brooks et al., 2011; Stan Development Team, 2014; Stan Development Team, 2016). The HMC algorithm commences at a specified initial set of parameters. These parameters include step size (for discretization time) and a number of steps (for total simulation time) whereby the simulation are mainly sensitive to the setting of these parameters (Stan Development Team, 2016). For the sake of maintaining the balance between the selections of the abovementioned two parameters, the No-U-Turn sampler (NUTS) is employed.

No-U-Turn sampler (NUTS)

NUTS is the HMC sampler that use gradients to guide updates (Stan Development Team, 2016). The NUTS automatically selects an appropriate number of leapfrog in each iteration in order to allow the proposals to traverse the posterior without doing unnecessary work (Hoffman & Gelman, 2011; Hoffman & Gelman, 2014). The motivation is to maximize the expected squared jump distance at each step and avoid the random-walk behavior that arises in random-walk Metropolis or Gibbs samplers when there is correlation in the posterior (Brooks et al., 2011).

HMC sampling using leapfrog steps

In HMC simulation, sampling is based on simulating the Hamiltonian of a particle with a starting position equal to the current parameter values and an initial momentum that is generated randomly. While implementing HMC in practice, the Hamiltonian dynamics of the particle is simulated using the leapfrog integrator, which discretizes the smooth path of the particle into a number of small time steps called leapfrog steps (Stan Development Team, 2016).

Metropolis acceptance

Next a Metropolis acceptance step is applied and a decision is made whether to update to the new state or keep the existing state (Stan Development Team, 2016). The metropolis adjustment is based on comparing log probabilities, defined by Hamiltonian for this case, which is the sum of the potential negative log probability and kinetic energies. The probability of rejection is determined by the accuracy of the leapfrog approximation to the true trajectory of the parameters. The idea here is to balance the rate of rejection versus the number of leapfrog steps.

Model Estimation Monitoring

Successful implementation of FB models requires a number of factors including sample size selection, divergence and amount of energy employed in the simulation. In HMC simulation, different parameters can be monitored during simulation. These include sample, divergence, energy, tree-depth, and step-size information. The performance of individual parameters during estimation can also be viewed individually.

Divergence information

In HMC simulation, divergent refers to the number of leapfrog transitions with diverging error. Considering the fact that NUTS terminates at the first divergence, divergent value will hence be 0 or 1 for each iteration. Thus, the average value of the divergent over all iterations represent the proportion of iterations with diverging error. Two main attributes that cause divergent transitions in HMC simulations are the presence of bugs in the model code and numerical instability in the leapfrog integrator used to simulate the Hamiltonian evaluation. Thus, evaluating the presence of divergence during MCMC simulation help in identifying light tails and incomplete exploration of the target distribution. The existence of divergent errors in the simulation will be identified by errors noted prior to termination of the simulation.

Energy information

In HMC simulation, the energy is the value of the Hamiltonian at each sample. The energy diagnostics assist in identifying overly heavy tails that are also challenging for sampling. Ostensibly, the energy diagnostics quantifies the heaviness of the tails of the posterior distribution. The energy diagnostic plot shows the overlaid histograms of the marginal energy distribution and the first differenced distribution. The discrepancies between the corresponding distributions are what determines the existence of heavy tails.

Tree-depth information

In HMC, tree-depth refers to the depth of tree used by NUTS. Configuring NUTS involves putting defining the upper limit of the depth of the tree to be used during the evaluation of individual iteration. This is monitored using the maximum depth parameter.

Step-size information

The integrator step size used in the Hamiltonian simulation is referred to step size information in HMC simulation. Use of the HMC requires the application of a numerical integrator in a step-size (Stan Development Team, 2016). On the other hand, if the step-size is too large, the leapfrog will be inaccurate and too many proposals will be repeated. If the step-size is too small, too many small steps will be taken by the leapfrog integrator leading to long simulation times per interval. Thus, the goal is to balance the acceptance rate between these extremes.

Model Convergence Parameters

A Markov chain produces samples from the target distribution after converging to equilibrium. Different diagnostic measures are employed to monitor the convergence of the respective chains. One of the parameters that is used to monitor the convergence of the chain is through the use of the potential scale reduction statistics Rhat $(\hat{R})$, (Gelman and Rubin, 1992). This parameter compares the current chain with other randomly initialized chains.

Specifically, the $\hat{R}$ statistic measures the ratio of the average variance of samples within each chain to the variance of the pooled samples across chains (Stan Development Team, 2016). The convergence of the monitored chains is considered to have achieved convergence when this statistic is equal to or less than one (Stan Modeling Team, 2016). Other convergence parameters include the ratio of the Monte Carlo (MC) errors relative to the standard deviations, whereby the ratio of the estimates less than 0.05 indicates convergence.

Model Diagnostics using Posterior Predictive Checks

The idea behind posterior predictive checks is analyzing the model fit of the developed model i.e. the ability of the particular model in predicting the response variable. To generate the replicated response variable, crashes for this case, the posterior predictive distribution equation is employed (Equation 3.25). the posterior predictive check supplements conventional techniques such as the residual analysis (Li et al., 2008).

$$P(y_{rep}|y = \int p(y_{rep}|\theta)d\theta, p(y_{rep}|y) = \int p(y_{rep}|\theta)p(y_{rep}|y)d\theta \tag{3.25}$$

Discussion of the Results

Prior specification

To obtain the FB estimates of the unknown parameters, it is required to specify prior distributions for the hyper-parameters. The most commonly used priors are vague normal distributions (with zero mean and large variance) for the regression parameters. The posterior distributions needed in the FB method were sampled using the MCMC, specifically HMC algorithm in the R software. The posterior estimates of the model's parameters for the FB methods were obtained via four independent chains with 100,000 iterations where 50,000 were used as a burn-in sample. The model chains convergence was monitored using three main statistical measures, i.e. Rhat, ratio of Monte Carlo standard error to the posterior standard deviation, and autocorrelation plots for each of the four chains.

Posterior distribution and cross-validation results

The posterior means and the 95th percentile Bayesian credible intervals (BCIs) of the posterior distributions for each of the crash category are shown in Tables 3.3 through 3.7. The predictor variable is considered to be significant at 95% BCI if the values of the 2.5% and 97.5% percentiles do not include zero (0) i.e., they are both negative or they are both positive.

Total crashes posterior distribution

The results for the evaluation of total crashes model are shown in Table 3.3. According to the results, model 3 has the highest prediction accuracy with the lowest LOO cross-validation value of -2458.9 and the standard error (SE) of 42.9 as compared to model 2 (Elpd LOO = -2553.7, SE=58. 2), and model 1 (Elpd LOO = -2585.4, SE = 62.8) (see the bottom of Table 3.3). Thus, model 3 was selected for further evaluation of the safety effectiveness of PCS to drivers on treatment intersections. Three predictor variables; AADT on the major approach, speed limit on the major approach, and the number of lanes on the minor approach, were significant at 95% confidence level. All the three variables are noted to be significantly positive. The coefficient of the AADT on the major approach (Mean = 0.345, 95% BCI (0.060, 0.132)) indicates that a prior increase in this variable increases the propensity of the overall crash occurrence. This is expected as the increase in traffic volume, which is one of the major exposure variables, is accompanied by an increase in heterogeneity in driving behavior and thus increases the probability of crash occurrence. The same explanation can be given for the other two significant variables, i.e., posted speed on the major approach and the total number of lanes on the minor approach. The results indicate that higher posted speed limit on the major approach influences the increase in the rate of crash occurrence by (Mean = 0.564, 95% BCI (0.198, 0.930)). Further, the increase in the number of lanes on the minor approach influences the rate of increase in crash occurrence by a mean coefficient of 0.448 with the 95% BCI of (0.127, 0.773).

Table 3.3 Posterior distribution summaries for total crashes

Variable/parameter	Model 1					Model 2					Model 3				
	Mean	**MCSE**	**SD**	**2.5%**	**97.5%**	**Mean**	**MCSE**	**SD**	**2.5%**	**97.5%**	**Mean**	**MCSE**	**SD**	**2.5%**	**97.5%**
Intercept	-3.348	0.0283	1.394	-6.143	-0.678	-3.436	0.0325	1.392	-6.196	-0.723	**-3.633**	**0.0060**	**1.491**	**-6.575**	**-0.726**
Intervention date	-0.012	0.0004	0.055	-0.121	0.097	-0.065	0.0006	0.076	-0.216	0.086	-0.133	0.0003	0.110	-0.350	0.083
Treatment by date	-0.004	0.0003	0.044	-0.089	0.082	-0.014	0.0003	0.044	-0.102	0.072	-0.103	0.0001	0.058	-0.216	0.010
Treatment by time	-0.048	0.0002	0.031	-0.109	0.013	-0.070	0.0003	0.039	-0.146	0.004	-0.021	0.0001	0.044	-0.108	0.066
Study period	0.041	0.0001	0.018	-0.007	0.076	0.055	0.0002	0.022	-0.011	0.099	0.069	0.0001	0.032	-0.007	0.132
ln(major AADT)	0.332	0.0028	0.136	0.068	0.602	0.339	0.0032	0.136	0.076	0.608	**0.345**	**0.0006**	**0.146**	**0.060**	**0.633**
ln(minor AADT)	0.047	0.0014	0.075	-0.100	0.192	0.046	0.0017	0.075	-0.105	0.194	0.055	0.0003	0.081	-0.105	0.214
Major approach posted speed <=40 mph															
Major approach posted speed >40 mph	0.566	0.0042	0.182	0.205	0.922	0.563	0.0054	0.191	0.189	0.942	**0.564**	**0.0008**	**0.186**	**0.198**	**0.930**
Number of major lanes	0.146	0.0043	0.163	-0.177	0.469	0.149	0.0050	0.162	-0.173	0.461	0.135	0.0008	0.164	-0.188	0.458
Number of minor lanes	0.454	0.0048	0.163	0.141	0.777	0.447	0.0047	0.163	0.129	0.769	**0.448**	**0.0008**	**0.165**	**0.127**	**0.773**
Other land use															
Commercial land use	0.179	0.0048	0.160	-0.139	0.485	0.177	0.0047	0.159	-0.137	0.490	0.168	0.0008	0.160	-0.146	0.481
Comparison intersections															
treatment intersections	0.224	0.0047	0.174	-0.115	0.571	0.258	0.0048	0.171	-0.081	0.595	0.157	0.0008	0.177	-0.192	0.504
Jump parameter						0.110	0.0009	0.110	-0.106	0.326	0.151	0.0003	0.122	-0.088	0.389
Sigma[random paired]											0.124	0.0001	0.016	0.094	0.158
Sigma[site]	0.762	0.0019	0.100	0.586	0.980	0.755	0.0023	0.100	0.580	0.972	0.735	0.0004	0.101	0.560	0.955
Mean PPD	6.509	0.0008	0.117	6.281	6.742	6.510	0.0008	0.117	6.286	6.741	6.509	0.0003	0.117	6.281	6.741
log-posterior	-2622	0.2356	13	-2648	-2598	-2623	0.2666	13	-2650	-2599	-3011	0.1158	24	-3060	-2964
Elpd LOO	-2585.4					-2553.7					**-2458.9**				
SE	62.8					58.2					**42.9**				

Property damage only (PDO) crashes posterior distribution

It can be noted from Table 3.4 that model 3 has the smallest Elpd LOO value of -2183.5 (SE = 41.6) cross-validation value as compared to model 1 (Elpd LOO = -2237.8, SE=48. 2) and model 2 (Elpd LOO = -2240.0, SE=48. 7). This indicates that model 3 has a better prediction accuracy as compared to the other two models and will hence be selected for further estimation of the safety effectiveness of PCS to drivers, based on PDO crashes. One parameter, study period, is further noted to be significantly different from zero at the 95 % BCI. The regression coefficient for treatment by date parameter (the parameter accounting for the difference in the slope of log crash frequency of time between treatment and comparison intersections during the before and after periods) is significantly negative (Mean = -0.189, 95% BCI (-0.326, -0.050)) suggesting a decrease in PDO crash frequency after the installation of PCS at the treatment intersections.

Five predictor variables were observed to be significant at the 95% BCI on the PDO crashes posterior distribution. These are AADT on the major (Mean = 0.530, 95% BCI (0.259, 0.801)) and minor (Mean = 0.228, 95% BCI (0.071, 0.390)) approaches, posted speed on the major approach (Mean = 0.482, 95% BCI (0.071, 0.817)), number of lanes on the minor approach (Mean = 0.340, 95% BCI (0.049, 0.643)), and last but not least, commercial land use (Mean = 0.309, 95% BCI (0.033, 0.595)). All of these variables are significantly positive, suggesting that their increase results in more occurrence of PDO crashes.

Table 3.4 Posterior distribution summaries for PDO crashes

Variable/parameter	Model 1					Model 2					Model 3				
	Mean	**MCSE**	**SD**	**2.5%**	**97.5%**	**Mean**	**MCSE**	**SD**	**2.5%**	**97.5%**	**Mean**	**MCSE**	**SD**	**2.5%**	**97.5%**
Intercept	-8.579	0.032	1.375	-11.209	-5.825	-8.611	0.032	1.363	-11.379	-6.010	**-7.725**	**0.032**	**1.393**	**-10.517**	**-5.008**
Intervention date	0.006	0.002	0.068	-0.130	0.137	0.028	0.002	0.090	-0.145	0.198	0.051	0.003	0.131	-0.211	0.297
Treatment by date	-0.047	0.001	0.056	-0.156	0.062	-0.043	0.001	0.060	-0.158	0.073	**-0.189**	**0.002**	**0.071**	**-0.326**	**-0.050**
Treatment by time	-0.053	0.001	0.041	-0.137	0.024	-0.045	0.001	0.050	-0.147	0.055	0.048	0.001	0.057	-0.066	0.158
Study period	0.141	0.001	0.021	-0.099	0.183	0.136	0.001	0.026	-0.085	0.186	0.127	0.001	0.038	-0.055	0.201
ln(major AADT)	0.601	0.003	0.134	0.333	0.857	0.602	0.003	0.131	0.356	0.871	**0.530**	**0.003**	**0.140**	**0.259**	**0.801**
ln(minor AADT)	0.252	0.002	0.078	0.092	0.411	0.254	0.002	0.080	0.102	0.414	**0.228**	**0.002**	**0.080**	**0.071**	**0.390**
Major approach posted speed <=40 mph															
Major approach posted speed >40 mph	0.466	0.004	0.168	0.139	0.792	0.472	0.004	0.162	0.136	0.789	**0.482**	**0.004**	**0.166**	**0.166**	**0.817**
Number of major lanes	0.017	0.004	0.147	-0.278	0.315	0.021	0.004	0.145	-0.253	0.311	0.028	0.003	0.146	-0.265	0.316
Number of minor lanes	0.306	0.004	0.146	0.021	0.596	0.300	0.003	0.146	0.014	0.582	**0.340**	**0.004**	**0.151**	**0.049**	**0.643**
Other land use															
Commercial land use	0.291	0.003	0.144	0.007	0.571	0.291	0.003	0.136	0.018	0.558	**0.309**	**0.003**	**0.143**	**0.033**	**0.595**
Comparison intersections															
Treatment intersections	0.287	0.004	0.166	-0.042	0.610	0.273	0.004	0.172	-0.072	0.605	0.092	0.004	0.184	-0.261	0.448
Jump parameter						-0.043	0.003	0.133	-0.313	0.209	-0.051	0.003	0.152	-0.347	0.251
Sigma[random paired]											**0.158**	**0.001**	**0.023**	**0.114**	**0.205**
Sigma[site]	0.564	0.002	0.080	0.424	0.729	0.561	0.002	0.079	0.425	0.734	**0.536**	**0.002**	**0.080**	**0.396**	**0.709**
mean PPD	4.143	0.002	0.092	3.961	4.327	4.148	0.002	0.095	3.962	4.331	4.144	0.002	0.094	3.965	4.328
log-posterior	-2323	0	13	-2349	-2299	-2325	0.000	13	-2352	-2302	-2777	1	24	-2825	-2733
Elpd LOO	-2237.8					-2240.0					**-2183.5**				
SE	48.2					48.7					**41.6**				

Fatal and injury (FI) crashes posterior distribution

Because there were just a few fatal crashes during the study period, fatal and injury crashes were combined together and denoted as FI. Model 3 continue to have the best prediction accuracy, according to the cross-validation results of this crash category shown in Table 3.5 (bottom two rows). Hence further discussion is based on the results of model 3 only. According to Table 3.5, two parameters are observed to have regression coefficients which are significant at the 95% BCI. The negative sign of the intervention parameter (Mean = -0.299, 95% BCI (-0.457, -0.139)) indicates a general decrease in FI for the after period. The jump parameter (Mean = 0.385, 95% BCI (0.111, 0.646)) indicates that there is a sudden increase in FI crash frequency at treatment intersections. Further, three predictor variables are observed to be significant at the 95% BCI. These are AADT on the major approach (Mean = 0.509, 95% BCI 0.246, 0.769)), AADT on the minor approach (Mean = 0.226, 95% BCI (0.083, 0.371)), and the posted speed limit on the major approach (Mean = 0.435, 95% BCI (0.105, 0.756)). The positive sign of the coefficients of these regression coefficients indicates their direct influence on FI crash frequency. Ostensibly, crashes occurring at higher speed are prone to have increased severity level.

Table 3.5 Posterior distribution summaries for FI crashes

Variable/parameter	Model 1					Model 2					Model 3				
	Mean	**MCSE**	**SD**	**2.5%**	**97.5%**	**Mean**	**MCSE**	**SD**	**2.50%**	**97.5%**	**Mean**	**MCSE**	**SD**	**2.5%**	**97.5%**
Intercept	-6.460	0.0317	1.338	-9.115	-3.905	-6.567	0.030	1.298	-9.165	-4.023	**-6.603**	**0.030**	**1.333**	**-9.226**	**-3.937**
Intervention date	-0.162	0.0015	0.066	-0.287	-0.034	-0.298	0.002	0.083	-0.459	-0.134	**-0.299**	**0.002**	**0.082**	**-0.457**	**-0.139**
Treatment by date	0.177	0.0015	0.062	-0.057	0.301	0.147	0.001	0.065	-0.017	0.271	0.048	0.001	0.063	-0.024	0.270
Treatment by time	-0.064	0.0010	0.044	-0.151	0.022	-0.042	0.001	0.054	-0.249	-0.033	-0.043	0.001	0.051	-0.047	0.245
Study period	0.058	0.0005	0.021	0.019	0.098	0.093	0.001	0.025	-0.042	0.141	0.093	0.001	0.024	-0.046	0.139
ln(major AADT)	0.498	0.0031	0.133	0.235	0.767	0.508	0.003	0.130	0.248	0.758	**0.509**	**0.003**	**0.131**	**0.246**	**0.769**
ln(minor AADT)	0.228	0.0017	0.075	0.084	0.373	0.223	0.002	0.074	0.077	0.366	**0.226**	**0.002**	**0.073**	**0.083**	**0.371**
Major approach posted speed <=40 mph															
Major approach posted speed >40 mph	0.445	0.0043	0.166	0.113	0.774	0.451	0.004	0.164	0.133	0.774	**0.435**	**0.004**	**0.166**	**0.105**	**0.756**
Number of major lanes	0.093	0.0039	0.145	-0.185	0.384	0.084	0.003	0.141	-0.188	0.363	0.083	0.004	0.143	-0.189	0.382
Number of minor lanes	0.060	0.0035	0.144	-0.220	0.354	0.067	0.004	0.146	-0.225	0.355	0.088	0.004	0.148	-0.202	0.385
Other land use															
Commercial land use	0.163	0.0037	0.139	-0.090	0.442	0.148	0.003	0.139	-0.117	0.421	0.159	0.004	0.148	-0.120	0.463
Comparison intersections															
Treatment intersections	-0.392	0.0048	0.172	-0.714	-0.057	-0.285	0.004	0.175	-0.620	0.063	-0.275	0.004	0.161	-0.583	0.047
Jump parameter						0.383	0.003	0.143	0.107	0.654	**0.385**	**0.003**	**0.138**	**0.111**	**0.646**
Sigma[random paired]											**0.476**	**0.002**	**0.089**	**0.314**	**0.667**
Sigma[site]	0.551	0.0018	0.079	0.416	0.730	0.549	0.002	0.079	0.413	0.722	**0.081**	**0.002**	**0.066**	**0.000**	**0.236**
mean PPD	4.401	0.0022	0.097	4.220	4.591	4.403	0.002	0.094	4.217	4.584	4.400	0.002	0.097	4.211	4.592
log-posterior	-2385	0.2886	13	-2411	-2360	-2383	0.307	13	-2409	-2358	-2497	0.329	14	-2527	-2471
Elpd LOO	-2319.9					-2310.8					**-2264.1**				
SE	53.9					53.2					**45**				

Rear-end crashes posterior distribution

In this study, the posterior distribution for the rear-end crashes were also developed, and are summarized in Table 3.6. Using the LOO cross-validation, model 3 is identified to have the best fit, hence it will be used for further discussion and estimation of index of effectiveness. Based on the results shown in Table 3.6, the regression coefficient for the parameter intervention date is significantly negative at 95% BCI (Mean = -0.125, 95% BCI (-0.198, -0.052,)), suggesting the general decrease in rear-end crashes for the after period for comparison and treatment intersections. Further, three explanatory variables are also significant at the 95% BCI. These are AADT on the major street (Mean = 0.142, 95% BCI (0.328, 0.886)), posted speed on the major street (Mean = 0.516, 95% BCI (0.187, 0.844)), and number of lanes on minor street (Mean = 0.368, 95% BCI (0.068, 0.671)). Their increase would result in the higher occurrence of rear-end crashes, due to the positivity of their regression coefficients.

Table 3.6 Posterior distribution summaries for rear-end crashes

	Model 1					Model 2					Model 3				
Variable/parameter	**Mean**	**MCSE**	**SD**	**2.5%**	**97.5%**	**Mean**	**MCSE**	**SD**	**2.5%**	**97.5%**	**Mean**	**MCSE**	**SD**	**2.5%**	**97.5%**
Intercept	-7.842	0.019	1.412	-10.632	-5.096	-7.858	0.021	1.439	-10.688	-5.077	**-7.511**	**0.022**	**1.462**	**-10.382**	**-4.612**
Intervention date	-0.122	0.0002	0.025	-0.171	-0.074	-0.114	0.000	0.029	-0.169	-0.057	**-0.125**	**0.000**	**0.037**	**-0.198**	**-0.052**
Treatment by date	0.024	0.001	0.080	-0.135	0.181	0.058	0.001	0.098	-0.133	0.251	-0.016	0.001	0.130	-0.272	0.238
Treatment by time	-0.140	0.000	0.054	-0.246	-0.034	-0.120	0.000	0.065	-0.247	0.004	-0.077	0.000	0.070	-0.214	0.059
Study period	0.156	0.001	0.076	0.007	0.303	0.165	0.001	0.078	0.012	0.319	0.052	0.001	0.088	-0.120	0.227
ln(major AADT)	0.641	0.002	0.140	0.367	0.919	0.640	0.002	0.141	0.357	0.922	**0.609**	**0.002**	**0.142**	**0.328**	**0.886**
ln(minor AADT)	0.114	0.001	0.085	-0.049	0.282	0.118	0.001	0.085	-0.050	0.283	0.108	0.001	0.086	-0.060	0.275
Major approach posted speed <=40 mph															
Major approach posted speed >40 mph	0.496	0.002	0.169	0.167	0.832	0.495	0.003	0.171	0.159	0.831	**0.516**	**0.003**	**0.167**	**0.187**	**0.844**
Number of major lanes	0.111	0.002	0.148	-0.179	0.402	0.116	0.002	0.148	-0.174	0.411	0.109	0.002	0.148	-0.180	0.403
Number of minor lanes	0.366	0.002	0.155	0.060	0.668	0.361	0.003	0.154	0.063	0.667	**0.368**	**0.002**	**0.153**	**0.068**	**0.671**
Other land use															
Commercial land use	0.195	0.002	0.143	-0.089	0.477	0.200	0.003	0.143	-0.079	0.480	0.198	0.002	0.144	-0.081	0.481
Comparison intersections															
Treatment intersections	-0.124	0.002	0.186	-0.491	0.242	-0.153	0.003	0.193	-0.538	0.213	-0.255	0.003	0.199	-0.648	0.135
Jump parameter						-0.103	0.001	0.170	-0.435	0.229	0.019	0.001	0.182	-0.337	0.373
Sigma[random paired]											**0.105**	**0.000**	**0.022**	**0.067**	**0.153**
Sigma[site]	0.550	0.001	0.082	0.409	0.732	0.553	0.001	0.081	0.414	0.730	**0.531**	**0.001**	**0.081**	**0.388**	**0.706**
mean PPD	3.075	0.001	0.080	2.920	3.234	3.074	0.001	0.081	2.918	3.233	3.075	0.001	0.080	2.918	3.234
log-posterior	-1975	0.223	13	-2001	-1950	-1976	0.217	13	-2002	-1952	-2536	0.361	24	-2584	-2489
Elpd LOO	-1871					-1869.2					**-1830.3**				
SE	39.3					39.2					**33.6**				

Angle crashes posterior distribution

The results of the posterior distributions of explanatory variables used in the Poisson-lognormal model for angle crashes is presented in Table 3.7. Using the LOO cross-validation, model 3 is identified to have the best fit, hence it will be used for further discussion and estimation of index of effectiveness. From Table 3.7, it can be observed that the regression coefficients for three explanatory variables are significant at the 95% BCI. These are AADT on the major street (Mean = 0.142, 95% BCI (0.328, 0.886)), posted speed on the major street (Mean = 0.516, 95% BCI (0.187, 0.844)), and number of lanes on the minor street (Mean = 0.368, 95% BCI (0.068, 0.671)). Because the coefficients of these three significant model variables are positive, the results suggest that the increase in their values would cause the increase in angle crashes.

Table 3.7 Posterior distribution summaries for angle crashes

Variable/ parameter	Model 1					Model 2					Model 3				
	Mean	**MCSE**	**SD**	**2.5%**	**97.5%**	**Mean**	**MCSE**	**SD**	**2.5%**	**97.5%**	**Mean**	**MCSE**	**SD**	**2.5%**	**97.5%**
Intercept	-8.934	0.023	1.584	-12.013	-5.846	-9.080	0.021	1.551	-12.124	-6.113	**-8.859**	**0.0077**	**1.608**	**-12.008**	**-5.699**
Intervention date	0.033	0.0002	0.028	-0.023	0.087	0.091	0.0003	0.037	-0.020	0.164	0.103	0.0001	0.044	-0.017	0.190
Treatment by date	0.061	0.001	0.087	-0.108	0.232	-0.167	0.001	0.125	-0.411	0.078	-0.191	0.0004	0.152	-0.490	0.106
Treatment by time	0.071	0.0004	0.052	-0.031	0.173	-0.027	0.0005	0.065	-0.154	0.102	-0.016	0.0002	0.071	-0.155	0.124
Study period	-0.004	0.0005	0.070	-0.139	0.133	-0.025	0.0005	0.070	-0.162	0.112	-0.054	0.0002	0.084	-0.219	0.110
ln(major AADT)	0.602	0.002	0.155	0.303	0.908	0.609	0.002	0.154	0.311	0.911	**0.572**	**0.0008**	**0.159**	**0.260**	**0.880**
ln(minor AADT)	0.269	0.001	0.086	0.100	0.438	0.267	0.001	0.087	0.100	0.440	**0.275**	**0.0004**	**0.091**	**0.096**	**0.453**
Major approach posted speed <=40 mph															
Major approach posted speed >40 mph	0.351	0.003	0.188	-0.012	0.724	0.354	0.003	0.190	-0.016	0.728	0.360	0.0009	0.190	-0.009	0.736
Number of major lanes	0.151	0.002	0.161	-0.162	0.467	0.154	0.002	0.164	-0.165	0.483	0.148	0.0009	0.164	-0.173	0.469
Number of minor lanes	0.144	0.003	0.166	-0.180	0.472	0.145	0.003	0.168	-0.188	0.473	0.158	0.0009	0.167	-0.169	0.488
Other land use															
Commercial land use	0.308	0.002	0.158	0.001	0.620	0.307	0.003	0.161	-0.009	0.624	0.312	0.0008	0.161	-0.002	0.630
Comparison intersections															
Treatment intersections	-0.054	0.003	0.198	-0.436	0.335	0.089	0.003	0.206	-0.313	0.498	0.078	0.0010	0.214	-0.341	0.497
Jump parameter						0.439	0.001	0.174	0.092	0.779	**0.451**	**0.0004**	**0.186**	**0.086**	**0.816**
Sigma[random paired]											**0.124**	**0.0001**	**0.024**	**0.081**	**0.176**
Sigma[site]	0.682	0.002	0.105	0.503	0.912	0.685	0.001	0.106	0.504	0.918	**0.660**	**0.0005**	**0.105**	**0.479**	**0.892**
mean PPD	2.530	0.001	0.074	2.386	2.676	2.530	0.001	0.073	2.389	2.673	2.530	0.0002	0.073	2.388	2.675
log-posterior	-1882	0.224	13	-1909	-1858	-1880	0.211	13	-1906	-1856	-2433	0.112	24	-2481	-2387
Elpd LOO	-1747.9					-1752.6					**-1726.6**				
SE	42.2					41.8					**35.3**				

PCS Index of Effectiveness

The PCS treatment effectiveness results are presented in Table 3.8. The table summarizes the overall treatment effectiveness for total, PDO, FI, rear-end, and angle crashes and the associated 95% credible intervals. The index of effectiveness is considered significant at the 95 % BCI when the values of the 2.5% and 97.5% percentiles do not include zero (1) i.e., they are both less than one or they are both greater than one. The FB method here accounts for the regressions-to-the-mean (RTM) phenomenon by estimating the expected number of crashes on the treatment intersections had the PCSs not installed by using the posterior distribution of a number of crashes which combines the observed data with the prior information. Ostensibly, this approach recognizes that the observed crash count at a site from any one year is a noisy measurement of the true long-run mean crash frequency (Park et al., 2010).

In this study, three Poisson lognormal models were developed for each crash category included in the analysis. The model with the best prediction accuracy selected based on Elpd LOO was hence selected for further estimation of the index of effectiveness. It can be depicted that three crash types, i.e. total (Mean = 0.894, 95% BCI (0.828, 0.911)), PDO (Mean = 0.908, 95% BCI (0.838, 0.953)), and rear-end (Mean = 0.920, 95% BCI (0.842, 0.942)) crashes are significant at the 95 % BCI as they're significantly different from one. It is worth noting the index values obtained in this study are comparable to the findings obtained using the findings using the EB methodology as illustrated in chapter 2 of this thesis. Based on these findings, it can be concluded that PCSs improve the safety of drivers at signalized intersections. It can also be observed from Table 3.8 that the results of two crash categories such as FI (Mean = 0.957, 95% BCI (0.886, 1. 020)) and angle (Mean = 0.969, 95% BCI (0.931, 1.022)) crashes are less than one but are not significant at the 95 % BCI. It is worth noting the significant reduction in rear-end crashes was expected, as the cues offered by PCS to drivers prompt drivers on the amount of time remaining to clear the intersection, thus facilitating smooth deceleration for drivers as they approach the intersection prior to termination of their green phase.

Table 3.8 Safety effectiveness indices (CMF) for different crash categories

Crash type	CMF		
		BCI	
	Mean	2.5%	97.5%
Total crashes	**0.894**	**0.828**	**0.911**
PDO	**0.908**	**0.838**	**0.953**
FI	0.957	0.886	1.020
Rear-end	**0.920**	**0.842**	**0.942**
Angle	0.969	0.931	1.022

Further, taking an advantage of the benefit of the FB approach, the CMF referred here as the safety effectiveness index was obtained per each year for after installing the treatment, thus obtaining a trend on the effect of the PCS in yearly basis. Figures 3.2 and 3.3 shows a similar declining trend of the treatment effectiveness index for total and PDO crashes. From this trend, it may be concluded that the safety improvements of PCS to drivers improves in time after they are being installed at the respective intersections. However, the treatment effectiveness on rear-end crashes is observed to decline with post-treatment time, although the mean value is still less than one for all the three years. This situation may be influenced by an increase in the proportion of drivers who wants to clear the intersection before the termination of the green phase, while other drivers apply breaks to avoid being caught up at the intersection after the termination of the green phase.

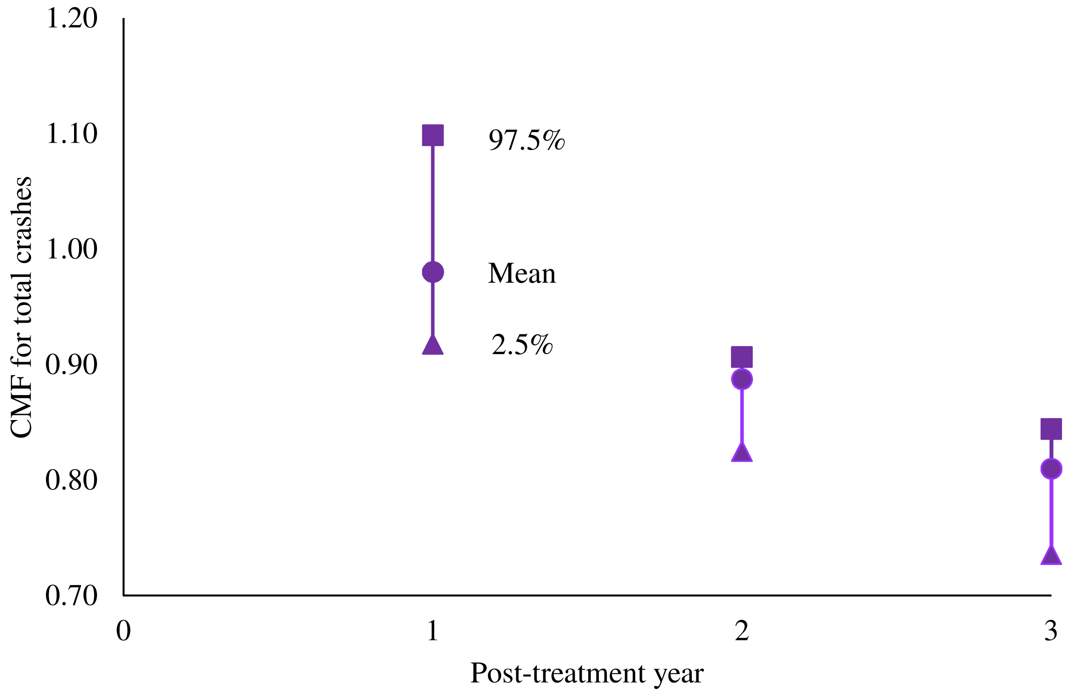


Figure 3.2. Yearly CMF for total crashes

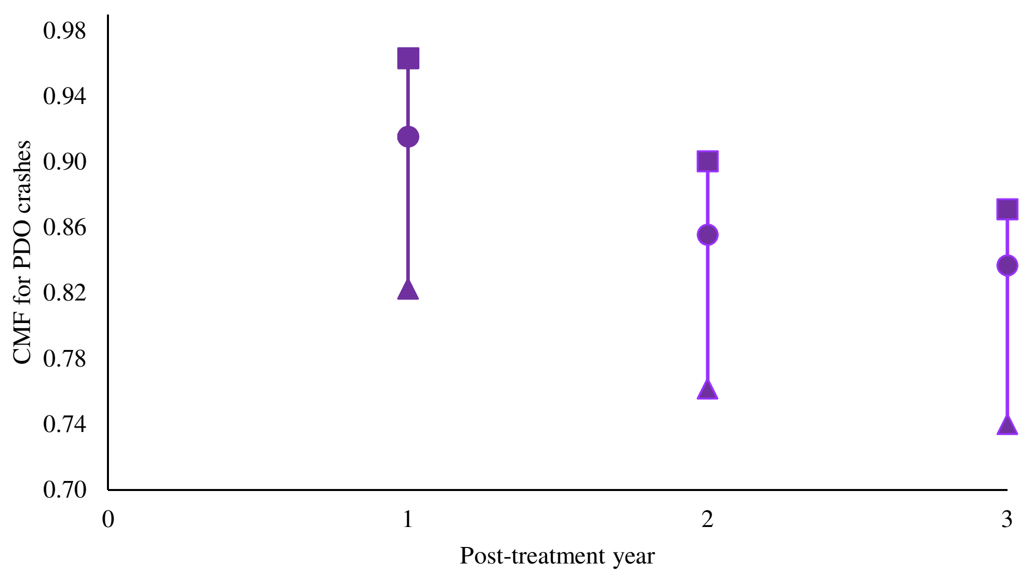


Figure 3.3. Yearly CMF for PDO crashes

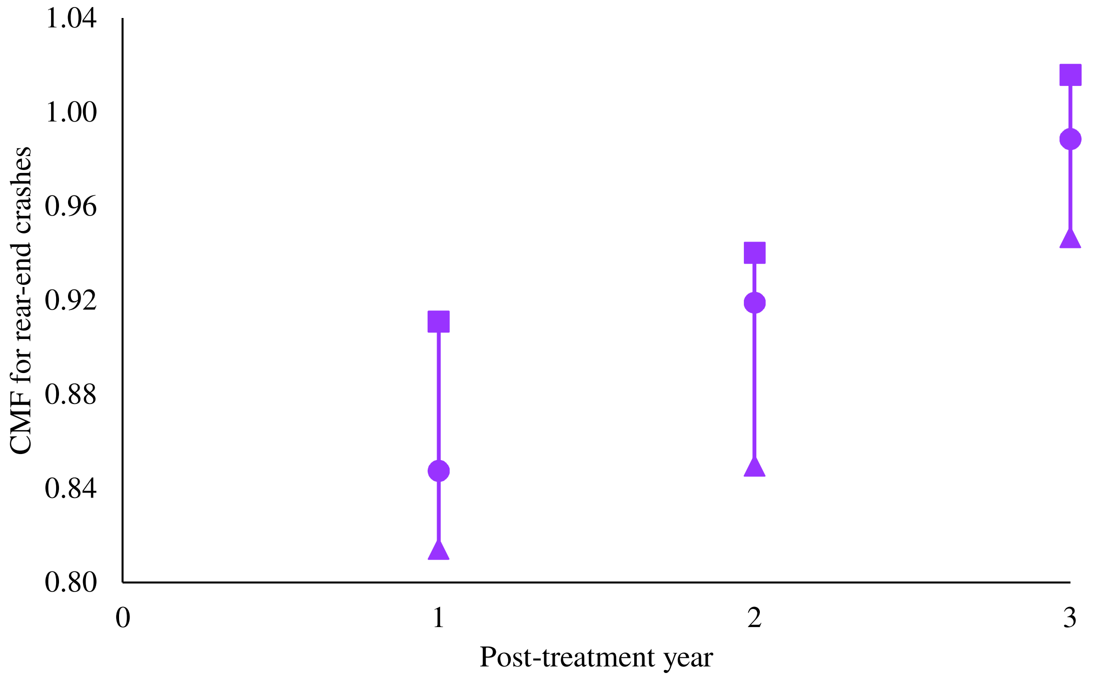


Figure 3.4. Rear-end CMF for rear-end crashes

Crash Modification Functions

CMFs are commonly used to evaluate the effects of the installed safety treatments. These safety indices are usually reported in terms of an average of the safety effect for a cohort of treatment intersections. Ostensibly, the estimated CMFs will vary based on individual site characteristics, mainly because they are exposed to different crash causation exposures i.e. traffic volume among other exposures. The relationship between the treatment effectiveness index (CMF) and a number of variables or a combination of variables is referred to as a crash modification function (CMFunction) as reported by different studies (Osama et al., 2016; Sacchi & Sayed, 2014b; Sacchi et al., 2015). Apart from crash exposure variables, the estimated indices of effectiveness are expected to vary with post-treatment time, i.e. the effect of the treatment is not instantaneous but rather spread over post-time periods. In this study, this scenario was achieved using the linear-intervention models.

In this study, traffic volume was employed as the main variable to incorporate into the CMFunctions for different crash types. The method used to develop the CMFunction is described in this section. Let $CMF_i = \theta_i$ denote an estimate of the index of effectiveness θ, where σ_i is a precision parameter for the respective index. A Gaussian linear regression model, presented in Equation 3.26 is used to model θ as the best estimate of the intervention effect across treatment intersections (Osama et al., 2016).

$$CMF_{i,y} = X_{i,y}\beta + \varepsilon_{i,y} \tag{3.26}$$

Where

$X_{i,y}\beta$ = linear predictor that represents the traffic volume covariate(s) at the intersection level that might influence the size of the intervention effect during the intervention, where y = 1, 2, and 3, for three years after installing PCSs.

β = vector of coefficients and a regression constant

$X_{i,y}$ = vector of covariates values for the ith intersection

$\varepsilon_{i,y}$ = error-term parameter across each intersection

Four different scenarios (Table 3.9) were fitted using the model in Equation 3.26. The four models were compared based on the LOOCV, where the model with the best fit was selected. It is worth noting that the results from the LOOCV for all models did not significantly differ, i.e. difference between scenarios <=0.1. Thus, the total entering traffic (AADT on the major approach + AADT on the minor approach) was used as the main explanatory variable for estimating the CMFunctions.

Table 3.9 Selection of predictor variable(s) in CMFunction formulation

Scenario	Independent variable
1.	X_i = AADT on the major approach
2.	X_i = AADT on the minor approach
3.	X_i = AADT on the major approach + AADT on the minor approach
4.	X_i = AADT on the major approach × AADT on the minor approach

To obtain Bayesian estimates of the unknown parameter, the R software was used. Four independent Markov chains were run using 10,000 iterations (the first 5,000 were discarded as burn-in runs). The model was fitted assuming the whole set of the regression parameters to be non-informative, i.e., they are normally distributed with a zero mean and a large variance. This procedure was repeated for the three crash categories with significant CMFs in this study. Once the best model was selected, the CMFunctions for each of the respective crash category and after-intervention year were estimated as summarized in Table 3.9.

Table 3.10 Posterior distribution summaries for the CMFunctions

Crash type	Variable/ parameter	After-intervention year								
		1			2			3		
		Mean	2.5%	97.5%	Mean	2.5%	97.5%	Mean	2.5%	97.5%
Total	Intercept β_0	0.574	0.273	0.877	0.582	0.332	0.833	0.721	0.546	0.895
	Total entering traffic β_1	0.039	0.01	0.068	0.029	0.005	0.053	0.008	-0.008	0.025
	Error-term ε_i	0.063	0.054	0.074	0.053	0.045	0.062	0.037	0.031	0.043
	Elpd_LOO (SE)	108.5 (4.9)			123.1 (7.3)			151.6 (7.3)		
PDO	Intercept β_0	0.33	-0.186	0.846	0.612	0.157	1.065	0.654	0.251	1.057
	Total entering traffic β_1	0.056	0.006	0.105	0.023	-0.020	0.067	0.017	-0.021	0.056
	Error-term ε_i	0.108	0.092	0.127	0.096	0.082	0.112	0.086	0.073	0.101
	Elpd_LOO (SE)	66.4 (8.3)			76.6 (11.1)			83.8 (6.8)		
Rear-end	Intercept β_0	0.276	-0.141	0.697	0.584	0.200	0.975	0.718	0.404	1.027
	Total entering traffic β_1	0.054	0.014	0.094	0.03	-0.008	0.066	0.024	-0.006	0.054
	Error-term ε_i	0.088	0.075	0.104	0.082	0.070	0.096	0.066	0.056	0.078
	Elpd_LOO (SE)	81.8 (5.3)			88.5 (9.6)			104.8 (6.1)		

Figures 3.5 through 3.7 show the plots of the obtained CMFunctions for total, PDO, and rear-end crashes, respectively. Each of the plots has a total of three post-treatment years estimated data. From these plots, it can be observed that the index of effectiveness (CMF) values generally increases with the traffic volume and then reaches an asymptotic value which is often close to one. This indicates less crash reduction for intersections with high traffic volume. In addition, the CMF values based on post-treatment time decreased generally over time in total crashes (Figure 3.5) and PDO crashes (Figure 3.6), indicating a consequent improvement of safety over time, but have increased for rear-end crashes (Figure 3.7).

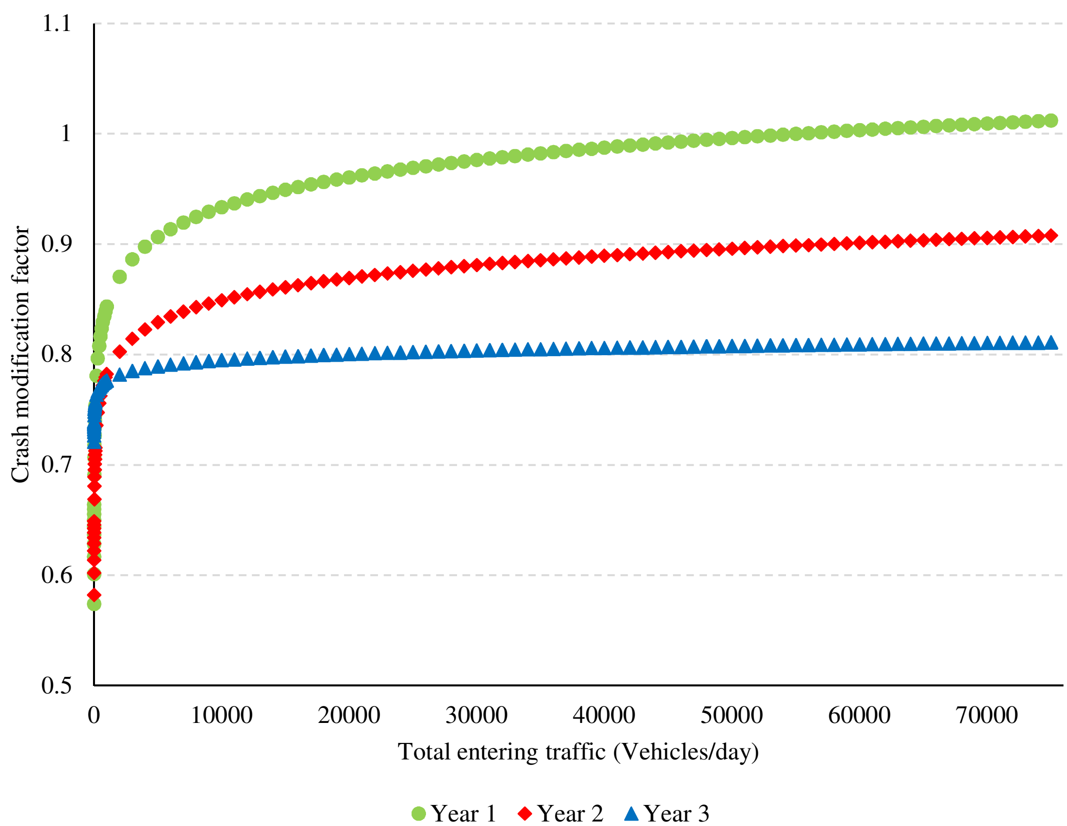


Figure 3.5. Total crashes CMFunction

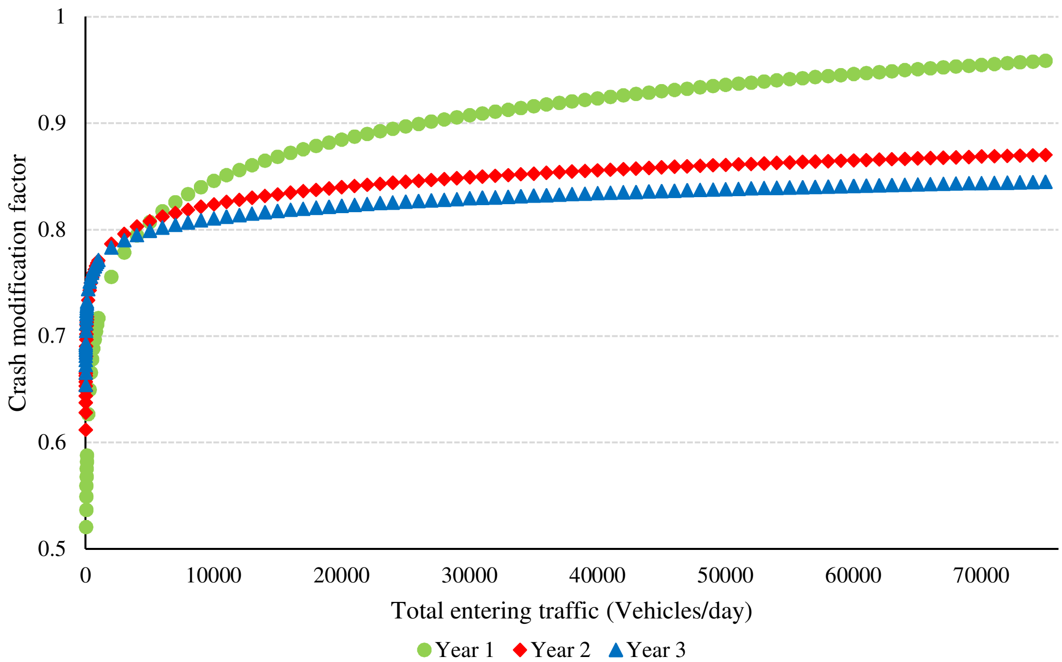


Figure 3.6. PDO CMFunction

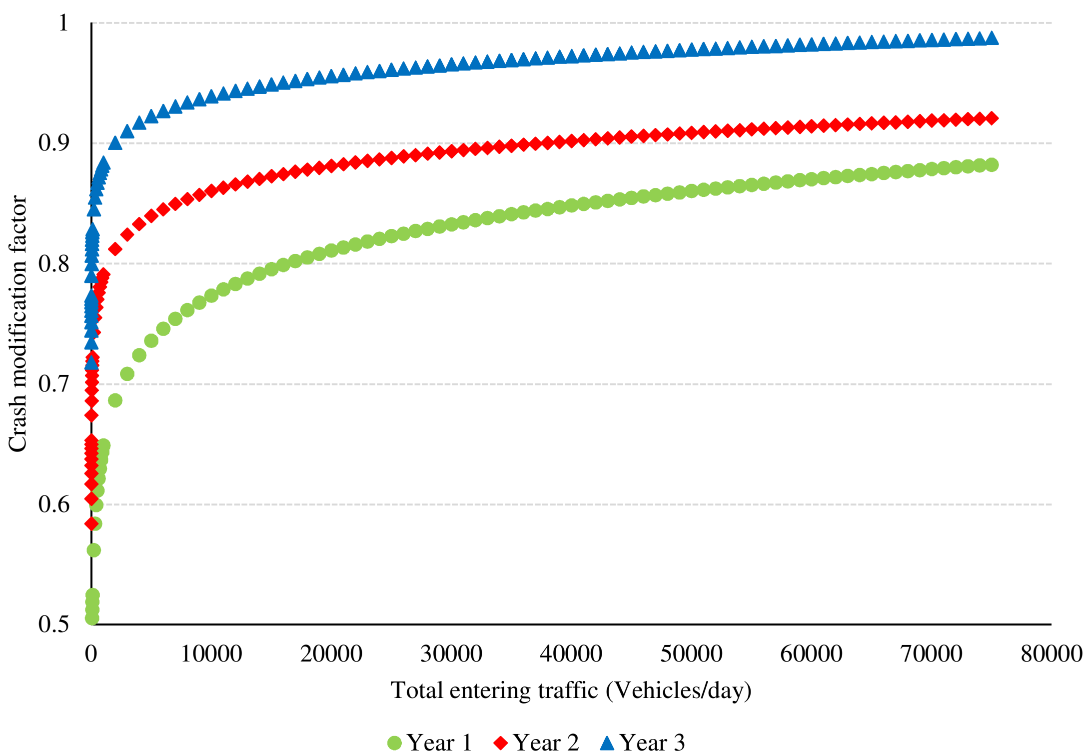


Figure 3.7. Rear-end CMFunction

Conclusions and Suggestions for Future Research

The use of full Bayes (FB) method is emerging as a more reliable approach in safety effectiveness studies that can account for the limitations of the conventional empirical Bayes (EB) approach. These EB method limitations include the substantial reliance on the safety performance functions (SPFs), which may result in the significant underestimation of the uncertainties in the crash modification factors (CMFs) estimated using the EB method. In this study, the FB approach was employed to investigate the safety effects of pedestrian countdown signals (PCSs) to drivers using the FB before-and-after with the comparison group approach. This manuscript has demonstrated the estimation of the safety effectiveness index using the FB method that employed the Poisson-lognormal model, with an intervention attribute, jump parameter, and additional random parameters. For the FB approach, uncertainties in the covariates are propagated throughout the model and are carried through the final safety effectiveness indices estimate.

Crash modification factors (CMFs) and crash modification functions (CMFunctions) were developed based on total crashes and two injury severity categories, i.e., property damage only (PDO) and fatal and injury (FI) crashes. Treatment effectiveness indices were also developed for two crash types – rear-end and angle crashes.

Based on the findings of this study, using the CMF as the safety effectiveness index, PCSs reduce total crashes by about 10% (Mean = 0.894, 95% BCI (0.828, 0.911)). Installation of PCSs is also observed to reduce PDO crashes by just less than 10% (Mean = 0.908, 95% BCI (0.838, 0.953)), according to the results of this study. Also, a reduction of 7.8% for rear-end (Mean = 0.920, 95% BCI (0.842, 0.942)) is observed after installing PCSs. Data do not suggest any significant effect of PCSs for angle and FI crashes.

The CMFunctions developed in this study indicate that the treatment effectiveness varies with post-intervention time and the crash exposure variable used in this study (traffic volume). Moreover, the results indicate the treatment effectiveness to increase over time for the post-treatment years for a total and PDO crashes. Conversely, the CMF values for rear-end crashes was observed to increase over time indicating a consequent decline in the improvement of safety over time.

Despite the advantages of the FB approach discussed herein, there is still room for improvement. Future studies can consider using non-linear time trend models for estimating the countermeasure safety effectiveness index. Further, more explanatory variables are proposed to be included in the CMFunctions estimation. This step can lead to more representative functions that better account for more variables that influence the variation of crash frequency. It will also be interesting to develop the CMF using the multivariate approach, especially in crash categories based on injury severities. Considering the fact that the presence of PCSs at signalized intersections may not only be beneficial to pedestrians, but also to drivers, it will be interesting to conduct a much broader study to confirm the findings of this study.

CHAPTER 4: CONCLUSIONS AND RECOMMENDATIONS

Overview

This thesis evaluates the safety effects of PCSs to drivers at signalized intersections using two methods - empirical Bayes (EB) followed by full Bayes (FB). The study employs the before-and-after study approach with the comparison group. For each site, six-year crash data (from years 2003 through 2014) are used – three years for the before- and the other three for the after- period. The analysis includes a total of 110 treatment and 93 comparison intersections. Comparison intersections are included to account for the external attributes that may influence the change in crash frequency apart from the main intervention, which is PCSs for this study. The study sites are located in the FDOT District 2, two cities in particular – Jacksonville and Gainesville.

The safety performance functions (SPFs) for the first method, i.e., EB approach were developed using the negative binomial model. In the case of the FB method, the Poisson-lognormal model with a piecewise linear change point for the before-and-after period was used. Several parameters including the site-level effect, jump, and paired-random effect parameters were introduced in the FB approach to improve model reliability. It is worth noting that the two count models used this study take into account the dispersion nature of crash occurrence.

Both of the count models used in this thesis incorporate a number of explanatory variables, including traffic volume and number of lanes on the major and minor approaches. They also incorporate the posted speed on the major approach and land use information. It is worth noting that the coefficients of each of these variables are positive for all the models, at different significance levels, indicating that the increase of these variables results in the increase the crash occurrence rates. It can also be inferred from the results that the coefficient of the traffic volume for the major approach is significantly positive and higher than the coefficient of the traffic volume of the minor approach, which is also positive and significant at 95% CI/BCI. This indicates that the influence of traffic volume on the major approach to increase in crash occurrence is higher than that of traffic volume on the minor approach.

Measures of Safety Effectiveness Index

The safety effectiveness of PCSs was measured using the safety effectiveness index known as the Crash Modification Factor (CMF). CMFs were developed for various crash categories based on type and injury severity. In addition to CMFs, Crash Modification Functions (CMFunctions) were developed. CMFunctions are mathematical expressions that relate the effectiveness of CMFs with varying exposure factors. For this study, traffic volume, in terms of AADT, was used as an exposure variable. CMFs and CMFunctions findings are summarized next.

Summary of Findings for Crash Modification Factors

Table 4.1 summarizes the computed CMFs estimated for different crash categories using the EB and FB approaches. The results are based on the 95% CI for the EB approach and 95% BCI for the FB methodology. Reported in Table 4.1 are also the CRFs (Crash reduction factors), which are computed as 1 – CMF.

Overall, the results indicate that the installation of PCSs significantly reduced total collisions by about 10%, PDO by nearly 9%, and rear-end crashes by 8%. The significant reduction in rear-end collisions was expected, as the cues offered by PCS timers are likely to affect the decisions of drivers as they approach the intersection. Generally, there are no big differences between the CMF results of the EB and FB methods. The same observation has been reported by a number of previous studies, including Persaud et al. (2010) and Park et al. (2016).

Table 4.1 Treatment safety effectiveness index for different crash categories using the EB and FB approaches

Crash type	Safety effectiveness indexes from the EB method				Safety effectiveness indexes from the FB method			
	Mean CMF	CRF (%)	95% CMF CI		Mean CMF	CRF (%)	95% CMF BCI	
			Lower	Upper			2.50%	97.50%
Total crashes	**0.912**	**8.8**	**0.855**	**0.969**	**0.894**	**10.6**	**0.828**	**0.911**
PDO	**0.929**	**7.1**	**0.862**	**0.996**	**0.908**	**9.2**	**0.838**	**0.953**
FI	0.952	4.8	0.797	1.107	0.957	4.3	0.886	1.020
Rear-end	**0.920**	**8.0**	**0.889**	**0.951**	**0.920**	**8.0**	**0.842**	**0.942**
Angle	0.954	4.6	0.797	1.111	0.969	3.1	0.931	1.022

Summary of Findings for Crash Modification Functions

This study also developed the CMFunctions for different crash types with total entering traffic as the primary explanatory variable. Regarding the extra benefits of the FB approach, the linear time trend was added into the CMFunction computations. Using plots, the trend of the treatment effectiveness indices was observed to diminish with increase in post-intervention duration. In summary, the CMFunctions developed in this study clearly show that the treatment effectiveness varies considerably with post-treatment time and crash exposure variables such as traffic volume. It is worth mentioning that the developed CMFs are based on local characteristics. For a broader use of these CMFs, calibration may be necessary.

Intuitive and Counterintuitive Findings

The findings from both the EB and FB methods results indicated that the presence of PCSs at signalized intersections improves drivers' safety. Furthermore, the treatment effectiveness was observed to improve over post-treatment years for the total and PDO crashes. In contrast, the safety effectiveness of PCSs at signalized intersections was found to decline with post-treatment years. Despite suffering a declining trend, the CMF values for all the post-treatment years are still less than one indicating safety improvement. The reduction in rear-end crash frequency upon installation of PCSs can be explained by the increase in the number of drivers who use the information offered by the countdown timer to slow down upon the termination of the green phase. In contrast, other drivers may use the same information to speed up to clear the intersection, hence avoid being caught up to wait for the next green phase. Therefore, the increase in rear-end crash frequency with post-treatment time may be influenced by the increment of the ratio of two types of drivers, i.e. those who speed up and those who slow-down as the PCSs' timer approaches zero.

Study Limitations and Recommendations for Future Work

There are several limitations of this study. Unavailability of traffic volume (AADT) data, particularly on non-state maintained roadways limited the number of sites selected for the study to state-maintained intersections only. Further, because the minor street roadways for some of the intersections are local roads, reliable traffic volume data could not be obtained. This, in turn, resulted in dropping of many sites from the analysis due to the incompleteness of data. In addition, the process of retrieving information on the installation dates of PCSs was long and tedious. This was due to the absence of a database with the dates of PCSs installation. Maintaining a database with the records of the installation dates for traffic control devices including PCSs is necessary to aid continued efforts in evaluating the effectiveness of such devices.

It is also worth noting that the variation between the treatment effectiveness index and the post-treatment time was assumed to be linear in this study, which may not invariably be the case. Thus, a more reliable approach such as the application of autoregressive models which do not assume the variation to be linear is proposed. Also, more explanatory variables could be explored, especially for estimating CMFunctions, in future work. Incorporating more post-treatment years in the estimation of the CMFunction is also significant to distinguish whether there will be a change in the trend of safety effectiveness. It will also be interesting to develop CMFs using the multivariate approach, especially on crash categories based on injury severities.

Looking at the ancillary behaviors of drivers at intersections with red light cameras it will be beneficial to integrate these effects while estimating the safety effectiveness of PCSs to drivers at signalized crossings. Also because this study analyzed only four-legged intersection, it will be important for future research to expand the analysis to cover three-legged signalized intersections. It is worth noting that, these sites were dropped from the analysis in this thesis due to the limitation of the sample size. In view of the fact that the average value of the LOOCV value was used as a measure of the prediction accuracy in this thesis, it will be beneficial to check the sensitivity of this value for each of the data used in the analysis.

The FB probabilistic approach used in this study is still an evolving area of research. In other research fields, there is a push to move from probabilistic Bayesian techniques to optimization based Bayesian methods due to their many advantages, including computational efficiencies and the speed of model convergence. Methods such as graphical models, including Bayesian networks and hidden Markov model among others, have a potential use in highway safety modeling. For future work, there is a need to pursue the possibility of employing such methods in safety effectiveness studies as they are known to have computational benefits and more optimized results.

References for Chapters 1 and 4

AASHTO. (2010). Highway Safety Manual 1st edition. Washington, D.C: Transportation Research Board of the National Academies.

Ahmed, M. M., Abdel-Aty, M. & Park, J., (2015). Evaluation of the Safety Effectiveness of the Conversion of Two-Lane Roadways to Four-Lane Divided Roadways: Bayesian Versus Empirical Bayes. *Transportation Research Record: Journal of the Transportation Research Board*, 2515, 41-49.

Chen, P., Pai, C., Jou, R., Saleh, W., & Kuo, M., (2015). Exploring motorcycle red-light violation in response to pedestrian green signal countdown device. *Accident Analysis and Prevention*, 75, 128-136.

Gross, F., Bhagwant, P., & Craig, L., (2010). A guide to developing quality crash modification factors. Vienna: Vanasse Hangen Brustlin, Inc (VHB).

Hauer, E. (1997). Observational Before-After Studies in Road Safety. London: Pergamon Publications.

Herbel, S., Laing, L., & McGovern, C. (2010). Highway Safety Improvement Program (HSIP) Manual FHWA-SA-09-029. Washington DC: Department of Transportation: Federal Highway Administration.

Huang, H., & Zeeger, C., (2000). The effects of pedestrian countdown signals in Lake Buena Vista. Chapel Hill: Highway Safety Research Center University of North Carolina.

Kimberly, E. A., Ruihua, T. & Mangum, B. C., (2004). Evaluation of Pedestrian Countdown Signals in Mongomery County, Maryland. *Transportation Research Record: Journal of the Transportation Research Board*, 1878, 36-41.

Carriquirry, A., Pawlovich, M., Welch, T., (2008). The choice of statistical models in road safety countermeasure effectiveness studies in Iowa. *Accident Analysis & Prevention*, 40(4), 1531–1542.

Nambisan, S. S., & Karkee, G. J., (2010). Do Pedestrian Countdown Signals Influence Vehicle Speeds?. *Transportation Research Record: Journal of the Transportation Research Board*, 2149, 70-76.

National Cooperative Highway Research Program (NCHRP). (2008). Report 617 Accident Modification Factors for Traffic Engineering and ITS Improvement. Washington D.C: Transportation Research Board of the National Academies.

Park, J., Abdel-Aty, M., & Lee, J. (2016). Use of empirical and full Bayes before–after approaches to estimate the safety effects of roadside barriers with different crash conditions. *Journal of Safety Research*, *58*, 31–40. https://doi.org/10.1016/j.jsr.2016.06.002

Sacchi, E. & Sayed, T., (2014). Accounting for heterogeneity among treatment sites and time trends in developing crash modification functions. *Accident Analysis & Prevention*, 72, 116–126.

Schmitz, N. J .(2011). The effects of pedestrian countdown timers on safety and efficiency of operation of operations at signalized intersections. Lincoln, Nebraska: University of Nebraska.

Scott, A. C., Atkins, K. N., Bentzen, B. L., & Barlow, J. M., (2012). Perception of Pedestrian Traffic Signals by Pedestrians with Varying Levels of Vision. *Transportation Research Record: Journal of the Transportation Research Board*, 2299, 57-64.

University of North Carolina Highway Safety Research Center (HSRC). (2013). CMF most wanted list. ⟨http://www.cmfclearinghouse.org/most_wanted.cfm⟩ (April 6, 2016).

U.S. Department of Transportation Federal Highway Administration. (2003). Manual on Uniform Traffic Control Devices for Streets and Highways (MUTCD). Washington D.C.

U.S. Department of Transportation Federal Highway Administration. (2009). Manual on Uniform Traffic Control Devices for Streets and Highways (MUTCD). Washington D.C.

Vasudevan, V., Pulugurtha, S. S., Nambisan, S. S., & Dangeti, M. R., (2011). Effectiveness of Signal-Based Countermeasures for Pedestrian Safety: Findings from Pilot Study. *Transportation Research Record: Journal of the Transportation Research Board*, 2264, 44-53.

References for Chapter 2/Paper 1

AASHTO. (2010). Highway Safety Manual 1st edition. Washington, D.C: Transportation Research Board of the National Academies.

Agbelie, B. R. D. K. and Roshandeh, A. M. 2015. Impacts of Signal-Related Characteristics on Crash Frequency at Urban Signalized Intersections. *Journal of Transportation Safety and Security*, 7(3), pp. 199-207.

Chen, H., Liu, P., Lu, J. J., and Behzadi, B. 2009. "Evaluating the Safety Impacts of the Number and Arrangement of Lanes on Freeway Exit Ramps." *Accident Analysis & Prevention* 41 (3): 543–51. doi:10.1016/j.aap.2009.01.016.

Chen, P., Pai, C., Jou, R., Saleh, W., & Kuo, M., (2015). Exploring motorcycle red-light violation in response to pedestrian green signal countdown device. *Accident Analysis and Prevention*, 75, 128-136.

Chen, F., Chen, S. and Ma, X. 2016. Crash Frequency Modeling Using Real-Time Environmental and Traffic Data and UnBalanced Panel Data Models. *International Journal of Environmental Research and Public Health*, 13(6: 609).

Eccles, K. A., Tao, R., & Mangum, B. C., (2004). Evaluation of Pedestrian Countdown Signals in Mongomery County, Maryland. *Transportation Research Record: Journal of the Transportation Research Board*, 1878, 36-41.

Elekwachi, O. L. (2010). Empirical investigation of the effect of countdown pedestrian signals on driver behavior and capacity at urban signalized intersections. PhD Dissertation Morgan State University.

Florida Department of Transportation, FDOT., (2012). Florida Strategic Highway Safety Plan. Government, Tallahassee, Florida: Florida Department of Transportation.

Gross, F., Bhagwant, P., Craig, L., (2010). A guide to developing quality crash modification factors. Vienna: Vanasse Hangen Brustlin, Inc (VHB).

Hauer, E. (1997). Observational Before-After Studies in Road Safety. London: Pergamon Publications.

Hilbe, J. M. (2012). Negative binomial regression. Cambridge U.K : Cambridge University Press.

Huang, H., & Zeeger, C., (2000). The effects of pedestrian countdown signals in Lake Buena Vista. Chapel Hill: Highway Safety Research Center University of North Carolina.

Jacob, Benjamin. (2015). Top Pedestrian Crash Intersections and Segments by Florida FDOT District. March 17. Accessed March 30, 2017. https://www.arcgis.com/home/item.html?id=cd2b3929ecc048839c4cb08b2aaa4a2d.

Kwigizile, V., Oh, J., Houten, R. V., Prieto, D., Boateng, R., Rodriguez, L., Ceifetz, A., Yassin, J., Bagdade, J., & Andridge. P., (2015). Evaluation of Michigan's Engineering Improvements for Older Drivers. Lansing, Michigan: Michigan Department of Transportation: Research Administration.

Lambrianidou, P., Basbas, S., & Politis, I., (2013). Can pedestrians' crossing countdown signal timers promote green and safe mobility?. *Sustainable cities and Society*, 6, 33-39.

Long, K., Liu, Y., & Han, L. D. (2013). Impact of countdown timer on driving maneuvers after the yellow onset at signalized intersections: An empirical study in Changsha, China. Safety Science, 54, 8–16. https://doi.org/10.1016/j.ssci.2012.10.007

Markowitz, F., Sciortion, S., Fleck, J. L., & Yee, B. M., (2006). Pedestrian countdown signals: Experience with an extensive pilot installation. *Institute of Transportation Engineers*, 76, 43-48.

Nambisan, S. S., & Karkee, G. J., (2010). Do Pedestrian Countdown Signals Influence Vehicle Speeds?. *Transportation Research Record: Journal of the Transportation Research Board*, 2149, 70-76.

National Cooperative Highway Research Program (NCHRP). (2008). Report 617 Accident Modification Factors for Traffic Engineering and ITS Improvement. Washington D.C: Transportation Research Board of the National Academies.

Park, J., M. Abdel-Aty, J. Lee, and Lee, C., (2015). Developing crash modification functions to assess safety effects of adding bike lanes for urban arterials with different roadway and socio-economic characteristics. *Accident analysis and Prevention*, 74, 179-191.

Pulugurtha, S. S., Desai, A., & Pulugurtha, N. M., (2010). Are Pedestrian Countdown Signals Effective in Reducing Crashes?. *Traffic Injury Prevention*, 11 (6), 632-641.

University of North Carolina Highway Safety Research Center (HSRC). (2013). CMF most wanted list. ⟨http://www.cmfclearinghouse.org/most_wanted.cfm⟩ (April 6, 2016).

U.S. Department of Transportation Federal Highway Administration. (2003). Manual on Uniform Traffic Control Devices for Streets and Highways (MUTCD). Washington D.C.

U.S. Department of Transportation Federal Highway Administration. (2009). Manual on Uniform Traffic Control Devices for Streets and Highways (MUTCD). Washington D.C.

Schmitz, N. J .(2011). The effects of pedestrian countdown timers on safety and efficiency of operation of operations at signalized intersections. Phd Thesis, Lincoln, Nebraska: University of Nebraska.

Scott, A. C., Atkins, K. N., Bentzen, B. L., & Barlow, J. M., (2012). Perception of Pedestrian Traffic Signals by Pedestrians with Varying Levels of Vision. *Transportation Research Record: Journal of the Transportation Research Board*, 2299, 57-64.

Shen, Q. (2007). Development of Safety Performance Function for Empirical Bayes Estimation of Crash Reduction Factors. PhD Thesis, Miami: Florida International University.

Vasudevan, V., Pulugurtha, S. S., Nambisan, S. S., and Dangeti, M. R., (2011). Effectiveness of Signal-Based Countermeasures for Pedestrian Safety: Findings from Pilot Study. *Transportation Research Record: Journal of the Transportation Research Board*, 2264, 44-53.

Washington, P. S., Karlafris, G. M., & Mannering, L. F., (2003). Statistical and Econometric Methods for Transportation Data Analysis. Boca Raton London: Chapman & Hall/CRC.

References for Chapter 3/Paper 2

AASHTO, (2009). Highway Safety Manual 1st edition, Washington D.C.: Transportation Research Board of the National Academies.

Ahmed, M. M., Abdel-Aty, M. & Park, J., (2015). Evaluation of the Safety Effectiveness of the Conversion of Two-Lane Roadways to Four-Lane Divided Roadways: Bayesian Versus Empirical Bayes. *Transportation Research Record: Journal of the Transportation Research Board*, 2515, 41-49.

Akay, M. F., Abut, F., Daneshvar, S., & Heil, D. (2015). Prediction of Upper Body Power of Cross-Country Skiers Using Support Vector Machines. *Arabian Journal for Science and Engineering*, 40(4), 1045–1055. https://doi.org/10.1007/s13369-015-1588-y

Anastasopoulos, P.C. & Mannering, F.L., (2009). A note on modeling vehicle accident frequencies with random-parameters count models. *Accident Analysis & Prevention*, 41(1), 153–159.

Arlot, S. & Celisse, A., (2010). A survey of cross-validation procedures for model selection. *Statistics Survey*, 4, 40-79.

Aul, N. & Davis, G., (2006). Use of Propensity Score Matching Method and Hybrid Bayesian Method to Estimate Crash Modification Factors of Signal Installation. *Transportation Research Record: Journal of the Transportation Research Board*, 1950, 17-23.

Blincoe, L. J., Miller, T. R., Zaloshnja, E. & Lawrence, B. A., (2015). The Economic and Societal Impact Of Motor Vehicle Crashes, 2010 (Revised): DOT HS 812 013, Washington, DC: National Highway Traffic Safety Administration.

Brooks, S., Gelman, A., Jones, G. L. & Meng, X.-L., (2011). Handbook of Markov Chain Monte Carlo. 1st ed. Boca Raton: CRC Press.

Carriquiry, A. L. & Pawlovich, M., (2004). From Empirical Bayes to Full Bayes: Methods for Analyzing Traffic Safety Data, Iowa: Iowa State University.

Chen, Y. & Persaud, B., (2014). Methodology to develop crash modification functions for road safety treatments with fully specified and hierarchical models. *Accident Analysis and Prevention*, 70, 131–139. http://dx.doi.org/10.1016/j.aap.2014.03.012.

Chen, P., C. Pai, R. Jou, W. Saleh, and M. Kuo., (2015). Exploring motorcycle red-light violation in response to pedestrian green signal countdown device. *Accident Analysis and Prevention*, 75 , 128-136, doi:10.1016/j.aap.2014.11.016.

Davis, G. A. & Yang, S., (2001). Bayesian identification of high-risk intersections for older drivers via Gibbs sampling. *Transportation Research Record: Journal of the Transportation Research Board*, 1746, 84-89.

Elekwachi, O. L., (2010). Empirical investigation of the effect of countdown pedestrian signals on driver behavior and capacity at urban signalized intersections, Phd dissertation, Washington: Morgan State University.

El-Basyouny, K. & Sayed, T., (2011). A full Bayes multivariate intervention model with random parameters among matched pairs for before–after safety evaluation. *Accident Analysis & Prevention*, 43(1), 87–94.

El-Basyouny, K. & Sayed, T., (2009a). Accident prediction models with random corridor parameters. *Accident Analysis & Prevention*, 41(5), 1118–1123.

El-Basyouny, K. & Sayed, T., (2009b). Collision prediction models using multivariate Poisson-lognormal regression. *Accident Analysis & Prevention*, 41(4), 820–828.

Hauer, E., (2001). Overdispersion in modelling accidents on road sections and in Empirical Bayes estimation. *Accident Analysis & Prevention*, 33(6), 799–808.

Hauer, E., (2013). Safety Performance Functions: A workshop.

Hoffman-Jørgensen, J., (1994). Probability With a View Towards Statistics Volume 1. New York: Chapman & Hall.

Hoffman, M. D. & Gelman, A., (2011). The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo. arXiv:1111.4246v1.xi.

Hoffman, M. D. & Gelman, A., (2014). The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo. *Journal of Machine Learning Research*, 15, 1351-1381.

"HSRC", University of North Carolina: Highway Safety Research Center., (2013). Safety Performance Function Development Guide: Developing Jurisdiction-Specific SPFs, Chapel Hill, North Carolina: Highway Safety Research Center.

Kimberly, E. A., Ruihua, T. & Mangum, B. C., (2004). Evaluation of Pedestrian Countdown Signals in Mongomery County, Maryland. *Transportation Research Record: Journal of the Transportation Research Board*, 1878, 36-41.

Kitali, A. E., Sando, T., Castro, A., Kobelo, D., & Mwakalonge, J. (2017). Developing Crash Modification Factors to Quantify Impacts of Pedestrian Countdown Signals to Drivers. Presented at the 96th Annual Meeting of the Transportation Research Board. Washington D.C.

Kuhn, M. & Johnson, K., (2016). Applied Predictive Modeling. 5th ed. New York: Springer.

Lan, B., Persaud, B., Lyon, C., Bhim, R., (2009). Validation of a Full Bayes methodology for observational before–after road safety studies and application to evaluation of rural signal conversions. *Accident Analysis & Prevention*, 41(3), 574–580.

Li, S. C.-Y., Sayed, T. & El-Basyouny, K., (2013). Full Bayes Before-and-After Evaluation of Traffic Safety Improvements in City of Edmonton, Canada. *Transportation Research Record: Journal of the Transportation Research Board*, 2386, 189-194.

Li, W., Carriquirry, A., Pawlovich, M., Welch, T., (2008). The choice of statistical models in road safety countermeasure effectiveness studies in Iowa. *Accident Analysis & Prevention*, 40(4), 1531–1542.

Lu, J., (2013). Development of Safety Performance Functions for SafetyAnalyst Applications in Florida (doctoral dissertation), Miami: Florida International University.

Miaou, S. -P. & Lord, D., (2003). Modeling traffic crash-flow relationships for intersections: dispersion, functional form, and Bayes versus empirical Bayes. *Transportation Research Record: Journal of the Transportation Research Board*, 1840, 31-40.

Nambisan, S. S. & Karkee, G. J., (2010). Do Pedestrian Countdown Signals Influence Vehicle Speeds?. *Transportation Research Record: Journal of the Transportation Research Board* , 2149, 70-76.

NHTSA, (2010). Report to Congress NHTSA's Crash Data Collection Programs CongressNHTSA's Crash Data Collection Programs: DOT HS 811 337, Washington, DC: National Highway Traffic Safety Administration.

NHTSA, (2016). 2015 Motor Vehicle Crashes: Overview: DOT HS 812 318, Washington DC: National Highway Traffic Safety Administration.

Ntzoufras, I., (2009). Bayesian Modeling Using Winbugs. Athens: A John Wiley & Sons, Inc. Publication.

Osama, A., Sayed, T. & Sacchi, E., (2016). Crash Modification Functions for Installation of Left-Turn Lanes at Signalized Intersection Approaches. *Transportation Research Record: Journal of the Transportation Research Board*, 2583, 42–49. http://dx.doi.org/10.3141/2583-06.

Park, E. S. & Lord, D., (2007). Multivariate Poisson-lognormal models for jointly modeling crash frequency by severity. *Transportation Research Record: Journal of the Transportation Research Board*, 2019, 1-6.

Park, E.S., Park, J. & Lomax, T.J., (2010). A fully Bayesian multivariate approach to before–after safety evaluation. *Accident Analysis & Prevention*, 42(4), 1118–1127.

Park, J. et al., (2015). Developing crash modification functions to assess safety effects of adding bike lanes for urban arterials with different roadway and socio-economic characteristics. *Accident Analysis & Prevention*, 74, 179–191.

Pawlovich, M., Li, W., Carriquiry, A. & Welch, T., (2006). Iowa's Experience with Road Diet Measures: Use of Bayesian Approach to Assess Impacts on Crash Frequencies and Crash Rates. *Transportation Research Record: Journal of the Transportation Research Board*, 1953, 163-171.

Persaud, B. et al., (2010). Comparison of empirical Bayes and full Bayes approaches for before–after road safety evaluations. *Accident Analysis & Prevention*, 42(1), 38–43.

Persaud, B. & Lyon, C., (2007). Empirical Bayes before-after safety studies: Lessons learned from two decades of experience and future directions. *Accident Analysis and Prevention*, 39(3), 546–555.

Robert, C. P., n.d. The Metropolis–Hastings algorithm.

Sacchi, E. & Sayed, T., (2014). Accounting for heterogeneity among treatment sites and time trends in developing crash modification functions. *Accident Analysis & Prevention*, 72, 116–126.

Sacchi, E., Sayed, T. & Osama, A., (2015). Developing crash modification functions for pedestrian signal improvement. *Accident Analysis & Prevention*, 83, 47–56.

Saito, M., Brimley, B. K., & Schultz, G. G. (2011). Transportation Safety Data and Analysis Volume 2: Calibration of the Highway Safety Manual and Development of New Safety Performance Functions Report No. UT-10.12b . Salt Lake, Utah: Utah Department of Transportation-Research Division.

Sauber-Schatz, E. K., Ederer, D. J., Dellinger, A. M. & Baldwin, G. T., (2016). Vital Signs: Motor Vehicle Injury Prevention — United States and 19 Comparison Countries, Atlanta: Morbidity and Mortality Weekly Report (MMWR) Morb Mortal Wkly Rep 65 http://dx.doi.org/10.15585/mmwr.mm6526e1.

Schlüter, P. J., Deely, J. J. & Nicholson, A. J., (1997). Ranking and selecting motor vehicle accident sites by using a hierarchical Bayesian model. *The Statistician*, 46(3), 293-216.

Schmitz, N. J., (2011). The effects of pedestrian countdown timers on safety and efficiency of operation of operations at signalized intersections (doctoral dissertation), Lincoln, Nebraska: University of Nebraska.

Scott, A. C., Atkins, K. N., Bentzen, B. L. & Barlow, J. M., (2012). Perception of Pedestrian Traffic Signals by Pedestrians with Varying Levels of Vision. *Transportation Research Record: Journal of the Transportation Research Board*, 2299, 57-64. 10.3141/2299-07

Team, S. D., (2014). STAN: A C++ Library for Probability and Sampling, Version 2.5.0.

Team, S. D., (2016). Stan Modeling Language User's Guide and Reference Manual. Version 2.14.0.

Vasudevan, V., Pulugurtha, S. S., Nambisan, S. S. & Dangeti, M. R., (2011). Effectiveness of Signal-Based Countermeasures for Pedestrian Safety: Findings from Pilot Study. *Transportation Research Record: Journal of the Transportation Research Board*, 2264, 44-53. 10.3141/2264-06.

Vehtari, A., Gelman, A., & Gabry, J. (2016). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing*, 1–20. https://doi.org/10.1007/s11222-016-9696-4

Xie, K. et al., (2014). Crash frequency modeling for signalized intersections in a high-density urban road network. *Analytic Methods in Accident Research*, 2, 39–51.

Yang, H. et al., (2013). Modeling work zone crash frequency by quantifying measurement errors in work zone length. *Accident Analysis & Prevention*, 55, 192–201.

Zhou, M., Li, L., Dunson, D. & Carin, L., (2012). Lognormal and Gamma Mixed Negative Binomial Regression. *Proc Int Conf Mach Learn*, 2012, 1343-1350.

VITAE

ANGELA E. KITALI

Education

University of North Florida, Jacksonville, Florida

Masters of Science in Civil Engineering, August 2015-Present (4.0 GPA)

University of Dar es Salaam, Tanzania

Bachelor of Science in Civil and Transportation Engineering, July 2015 (3.5 GPA)

Relative Work Experience

University of North Florida, Jacksonville, Florida August 2015-Present

Graduate Research Assistant (Transportation safety modeling, operations, and microscopic traffic simulation studies)

1. Collecting, processing, interpreting, analyzing, and compiling traffic data obtained for research projects
2. Worked on a USDOT sponsored project "Evaluation of Pedestrian Signals to Reduce Intersection Crashes and Red Light Violations for Elderly Drivers"- Activities performed include creating crash modification factors, operational traffic data collection, and analysis
3. Evaluated the effect of driver age on vehicle deceleration rates using SHRP2 data and how it affects safety and travel time- activities performed includes traffic data collection, microsimulation of driver deceleration behaviors, surrogate safety analysis of conflicts, and propose improvements
4. Worked on a USDOT sponsored project "Micro-Analysis of Collisions in Crash Clusters: Creating Crash Patterns and Conducting a Driver Simulation Study"- Activities performed include a literature review on the subject in hand, and creating the crash pattern model
5. Reviewer of Journal of Advances in Transportation Studies (ATS) - Activities performed include reviewing and provide comments for three technical manuscripts

<u>Publications</u>

Published Papers

1. Evaluating aging pedestrian crash severity using Bayesian complementary log-log model for improved prediction accuracy-Journal of Transportation Research Board
2. Understanding the factors associated with severity of aging population-involved pedestrian crashes in Florida-Journal of Advances in Transportation Studies

Peer-reviewed papers conference proceedings

1. Using naturalistic driving data to analyze the effects of drivers' age deceleration on safety and travel time, a podium presentation done at the Fifth International Naturalistic Driving Research Symposium Blacksburg, Virginia from August 30-September 1, 2016
2. Developing crash modification factors to quantify impacts of pedestrian countdown signals to drivers. Transportation Research Board, Paper No. 17-05178, Washington D.C., January 2017
3. Impact of pedestrian countdown signals on vehicle approach speed for drivers of different age at signalized intersections. Transportation Research Board, Paper No. 17-06518, Washington D.C., January 2017
4. Evaluating aging pedestrian crash severity using Bayesian complementary log-log model for improved prediction accuracy. Transportation Research Board, Paper No. 17-06386, Washington D.C., January 2017
5. Simulation-based comparative performance measures for I-295 express lanes in Jacksonville, Florida. Transportation Research Board, Paper No. 17-06536, Washington D.C., January 2017
6. Operational Characteristics of the Newly Introduced Bus Rapid Transit in Dar es Salaam, Tanzania. Transportation Research Board, Paper No. 17-06618, Washington D.C., January 2017

Papers considered for journal publication

1. Appraisal of Safety Effects of Pedestrian Countdown Signals to Drivers Using Crash Modification Factors

2. Using naturalistic driving data to analyze the effects of drivers' age deceleration on safety and travel time- Journal of Safety Research

University of Dar es Salaam, Dar es Salaam Tanzania

1. Poster presentation on Tanzania women engineers convention exhibition (TAWECE) by IET titled "Strategy to improve the accessibility of public roads for people with disability in Tanzania"
2. Assessment of relationship between power demand and supply scenario to social-economic activities and GDP in Dar es Salaam

<u>Campus Involvement and Volunteer Experience</u>

1. Member of the Institute of Engineers Tanzania (IET)
2. National Society of black engineers, UNF chapter
3. American Society of Highway Engineers, UNF chapter
4. Volunteer at the Jacksonville 2015 Autonomous Vehicles Summit

<u>Awards and Funding</u>

1. Selected to attend and present at the Statewide Graduate Student Research Symposium-March 2017
2. Invitation to membership in the UNF Chapter of the Honor Society of **Phi Kappa Phi -**March 2017
3. Invitation to the UNF Chapter of **Phi Kappa Phi celebratory reception-**January 2017
4. UNF Chapter of **Phi Kappa Phi celebratory reception-**January 2016
5. 2015 Best Female Graduating Student Award, by Tanzania Engineers registration board, September 2015
6. Recipient of AASHTO Transportation Fellowship-August 2015
7. 2015 TAWECE conference 3rd winner of poster presentation titled "Strategy to improve the accessibility of public roads for people with disability in Tanzania"-July 2015
8. 2015 Powering Africa student research competition 1st winner, by US Embassy in collaboration with COSTECH-May 2015
9. Certificate of appreciation as an exhibitor during the CoET Undergraduate Students' Open Day- June 2014