5-1-2019

Modeling Driver-Pedestrian-Infrastructure Interactions at Signalized Midblock Crosswalks

Boniphace Kutela
kutela4@gmail.com

Repository Citation
Kutela, Boniphace, "Modeling Driver-Pedestrian-Infrastructure Interactions at Signalized Midblock Crosswalks" (2019). *UNLV Theses, Dissertations, Professional Papers, and Capstones*. 3635.
https://digitalscholarship.unlv.edu/thesesdissertations/3635

MODELING DRIVER-PEDESTRIAN-INFRASTRUCTURE INTERACTIONS AT SIGNALIZED MIDBLOCK CROSSWALKS

By

Boniphace Kutela

Bachelor of Science (Civil & Structural Engineering)
University of Dar es salaam
2010

Master of Engineering
Tennessee State University
2013

A dissertation submitted in partial fulfillment
of the requirements for the

Doctor of Philosophy - Civil and Environmental Engineering

Department of Civil and Environmental Engineering and Construction
Howard R. Hughes College of Engineering
The Graduate College

University of Nevada, Las Vegas
May 2019

UNLV | GRADUATE COLLEGE

Dissertation Approval

The Graduate College
The University of Nevada, Las Vegas

April 12, 2019

This dissertation prepared by

Boniphace Kutela

entitled

Modeling Driver-Pedestrian-Infrastructure Interactions at Signalized Midblock Crosswalks

is approved in partial fulfillment of the requirements for the degree of

Doctor of Philosophy - Civil and Environmental Engineering
Department of Civil and Environmental Engineering and Construction

Hualiang Teng, Ph.D.
Examination Committee Chair

Mohamed Kaseko, Ph.D.
Examination Committee Member

Ryan Sherman, Ph.D.
Examination Committee Member

Jin Ouk Choi, Ph.D.
Examination Committee Member

Mary Riddel, Ph.D.
Graduate College Faculty Representative

Kathryn Hausbeck Korgan, Ph.D.
Graduate College Dean

ABSTRACT

Cities and metropolitan areas are increasingly facilitating pedestrians' movement by the provision of pedestrian walking facilities. As pedestrian traffic increases, the risk of crash involvement increases, especially at midblock locations, where pedestrians are exposed to unsafe interactions with vehicular traffic. To improve pedestrians' safety at midblock locations, various countermeasures are provided, which include signalized crosswalks. Several studies have analyzed driver-pedestrian interactions, as well as pedestrian-infrastructure interactions at signalized midblock crosswalks. However, more in-depth studies are necessary, due to shortfalls of study assumptions, which have led to the application of improper statistical models, as seen in the literature. Improved models are crucial, as they can be used to evaluate the factors affecting the effectiveness of countermeasures at signalized midblock crosswalks. Moreover, there are several aspects of pedestrian-infrastructure interactions that have not been studied in the previous research. This study, therefore, attempts to improve the methodologies for analyzing driver-pedestrian-infrastructure interactions at signalized midblock crosswalks. Specifically, this study is aimed towards:

- Developing improved modeling methodology for the yielding compliance of drivers at signalized midblock crosswalks, which considers the time taken to yield right of way, and the transition states undergone during yielding.
- Analyzing the risks associated with driver-pedestrian interactions at signalized midblock crosswalks.
- Developing the framework for modeling the spatial and temporal crossing compliance of pedestrians at signalized midblock crosswalks.

- Evaluating the influence of various crosswalk features, such as signs and markings, traffic-related variables, and pedestrian related factors on the safe utilization of signalized midblock crosswalks; these include factors influencing drivers' yielding compliance, pedestrians' crossing compliance, and pedestrians' utilization of pushbuttons.

The study data were collected from a total of twenty signalized midblock crosswalks located in the Las Vegas, Nevada metropolitan area. These crosswalks have varying geometric configurations, signalizations, traffic characteristics, and pedestrian flows. Five types of signalization; Circular Flashing Beacons (CFBs), Circular Rapid Flashing Beacons (CRFBs), Rectangular Rapid Flashing Beacons (RRFBs), Pedestrian Hybrid Beacons (PHBs), and Traffic Control Signals (TCSs) were studied in this research. The observational survey method was applied for data collection, whereby video cameras were used to collect driver-pedestrian interactions. The data extraction was performed by reviewing the videos and recording the information of interest in a spreadsheet, with a total of 2638 pedestrians crossing incidents recorded for analysis. A descriptive analysis was performed, and several statistical models were developed.

Multistate hazard-based models are developed to model the yielding compliance of drivers. The transitional states while drivers are yielding right of way to pedestrians are defined as non-yield, "partial-yield" events (partial-yield, scenarios in which driver(s) in one lane yield, while other driver(s) in adjacent lane(s) in the same direction do not), and full-yield. Binary-based models are developed for modeling drivers' spatial yielding compliance, pedestrians' spatial crossing compliance, and pedestrians' temporal crossing compliance. Rare Events Logistic Regression (RELR) is applied to evaluate the occurrence of partial-yield events and near-miss

events. In addition to binary models, ordered models and multinomial models are developed and compared to model pedestrians’ spatiotemporal crossing compliance.

The results of the multistate models reveal that signal type, number of vehicles within effective crosswalk distance, yield-here sign, and crossing zone factors have similar influence for transition from non-yield to full-yield, non-yield to partial yield, and partial yield to full yield. Thus, the results of the binary models for yielding compliance are only partially comparable to one transition of the multistate model (non-yield to full yield). Through the Rare Event Logistic Regression (RELR) model, this study finds that near crash events are highly associated with a single cross stage, a high number of lanes, and night time. In addition, this study reveals that there is a strong association between partial-yield and near-miss events. Additionally, it is found that for every second that traffic continues to flow while pedestrians are waiting to cross, the probability of a partial-yield event occurring increases by 2.1%, while that of near-crash events increase by about 3%. Moreover, the influence of the crosswalk features and the distance at which drivers yield with respect to the yield line (spatial yielding) was assessed. The logistic regression results for associating drivers’ spatial yielding results shows that the odds for drivers’ spatial yielding are high if the crosswalks are equipped with Rectangular Rapid Flashing Beacons (RRFBs) at the advanced pedestrians crossing signs (APCSs), in the presence of “State Law” and “PED XING” signs. On the other hand, long distances from stripes to the yield lines, multiple cross stages, and high Annual Average Daily Traffic (AADT) are associated with decreased spatial yielding compliance.

Regarding pedestrian-infrastructure interactions, the logistic regression results reveal that the arrival sequence to a crosswalk has the highest impact on warning light activation tendencies. This means that the first arriving pedestrians are eight times more likely to press pushbuttons.

Moreover, males, the elderly, children, and teens are less likely to press pushbuttons. Furthermore, pedestrians who are involved in secondary activities, such as carrying/holding objects in their hands, have a relatively low odds ratio of pressing the pushbutton, while phone use is a statistically insignificant factor. Several infrastructure and traffic factors, including flash-based signal types (CRFBs, CFBs and RRFBs), a high number of lanes, residential land use, and higher oncoming vehicle speeds are associated with an increase of pushbutton pressing. Among the models applied for spatiotemporal crossing compliance, the logistic regression outperformed the multinomial logit and the ordered logit models. The logistic regression results reveal that the active WALK signal and a crossing incident involving female(s) only are the factors positively associated with pedestrians' spatiotemporal crossing compliance. On the other hand, wait time, children, and teens, as well as people who cross while using a phone or riding a bike are negatively associated with spatiotemporal crossing compliance.

Based on the study's findings, several recommendations are provided. The findings and recommendations from this study are expected to have academic, industry, and community benefits. Planners and engineers can benefit from this study by learning which countermeasures improve safety for both pedestrians and drivers. The models can be used by academicians and other practitioners to assess the scenarios in question. Improved pedestrian safety due to the selection of appropriate countermeasures, which fit a particular location, is a benefit that directly impacts the community.

ACKNOWLEDGEMENT

The journey to this Ph.D did not start five years ago, and it did not follow a straight path. First and foremost, I would like to thank God the almighty as it was a joke if you could have told me or any other person who knew me twenty years ago that one day, I will receive a Ph.D. But God had it planned. My sincere appreciations also go to my supervisor Professor Hualiang (Harry) Teng. For the past five years, the mentorship, financial support as well as moral supports that I have received from him are overwhelming. The committee members: Professor Mohamed Kaseko, Professor Ryan Sherman, Professor Jin Ouk Choi, and Professor Mary Riddel, have played a significant role in shaping this dissertation. My former supervisor during masters studies Dr. Deo Chimba is among the people that shaped me the way I am now. He has been very supportive on and off academia. My indebtedness also goes to my secondary school head master Mr. Chasilla as he was one of the earliest people to know that I will be someone in the future. I am also beholden to my friends in research, Mr. Emmanuel Kidando and Ms. Angel Edes. These two have motivated me to work hard every day to complete this dissertation. I would also like to thank my family. My wife Neema Langa has been there to encourage me during the times that I felt the dissertation was not moving forward. My kids Faith-Ariana Kutela and Brian Kutela, you are the reason I am working this hard, thank you for pushing me to work even harder. Lastly but not least, my parents Mr. & Mrs. Kutela for taking that first step to take me to school, and my cousin Edwin John who did not give up on me financially and morally during my secondary school education. I say, thank you.

DEDICATION

To my kids Faith-Ariana, Brian, and other coming,

take this as a challenge, I have set a bar you need to go beyond it.

To my wife Neema Langa

for your love, support, and understanding of my late nights absence when working on this dissertation.

To brothers, sisters, cousins, nieces, and nephews

if I did it you can do it too, never give up.

Most of all to my mom and dad,

for prayers, love, starting that step of taking me to school, and tireless efforts to make me fall in love with science subjects.

TABLE OF CONTENTS

CHAPTER 3: STUDY METHODOLOGY ... 49

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AADT	Annual Average Daily Traffic
AIC	Akaike Information Criterion
APCS	Advanced Pedestrian Crossing Sign
BIC	Bayesian Information Criterion
CCR	Crossing Compliance Rate
CFB	Circular Flashing Beacon
CRFB	Circular Rapid Flashing Beacon
FHWA	The Federal Highway Administration
LR	Logistic Regression
MUTCD	Manual on Uniform Traffic Control Devices
NHTSA	National Highway Traffic Safety Administration
NDOT	Nevada Department of Transportation
PHB	Pedestrian Hybrid Beacon
RELR	Rare Events Logistic Regression
RRFB	Rectangular Rapid Flashing Beacon
TCS	Traffic Control Signal

CHAPTER 1: INTRODUCTION

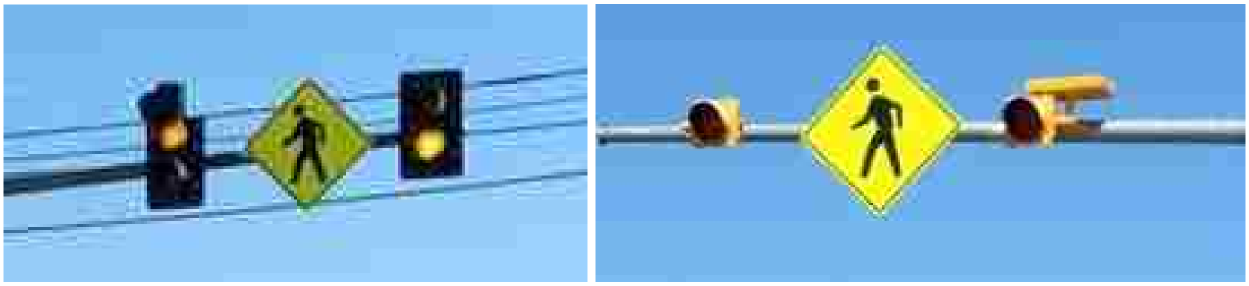
1.1. Motivational Background

Cities and metropolitan areas are increasingly facilitating pedestrians' movement through the provision of pedestrian walking facilities within downtown areas, school zones, recreational areas, or residential locations. With the increase in pedestrian traffic, the risk of crash involvement is increasing, especially when pedestrians are exposed to unsafe interactions with vehicular traffic. Pedestrians and bicyclists have higher odds of fatal crash involvement than other road users. Studies have shown that pedestrians are 1.5 times more likely to be involved in fatal crashes than vehicle occupants (Beck, Dellinger, and O 'neil 2007). It has been reported that most vehicle-pedestrian crashes occur at non-intersection locations (72%), in urban areas (76%), and when it is dark (74%) (NHTSA 2015). The most recent statistics show that the number of pedestrian fatalities in the United States has drastically increased compared to other traffic related fatalities. In fact, between 2007 and 2016, pedestrian fatalities increased by 27% compared to a 14% decrease in other traffic related fatalities (Retting 2017). The proportion of pedestrian fatalities to vehicular fatalities also increased, from 11% to 16% within the same period. This pedestrian fatality proportion record is the worst in the past 33 years (Retting 2017). At least 15 states had two fatalities per 100,000 people in 2016, which is double the number of states with similar fatality rates in 2014. The reasons most cited for vehicle-pedestrian crashes are visibility in darkness, failure to yield right of way, and improper crossing locations.

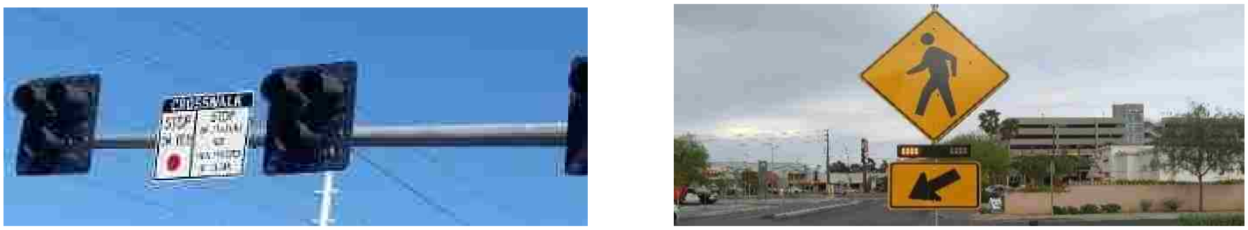
As most vehicle-pedestrian crashes occur at non-intersection locations, engineers and planners strive to provide dedicated crossing locations wherever there is a need for a substantial number of pedestrians to cross at non-intersection locations. These dedicated crossing locations are referred to as midblock crossings. The midblock crossings are accompanied by several treatments, including pedestrian signs, high-visibility markings, colored texture markings, refuge islands, in-pavement illumination, and pedestrian signals. The decisions for the types of crossing treatments applied are commonly based on the Manual on Uniform Traffic Control Devices (MUTCD) warranties (Manual on Uniform Traffic Control Devices (MUTCD) 2009). In addition, several states have established their own guidelines to suit their needs, by adopting "as is," modifying, or referencing the study by (Zegeer et al. 2005) as a source of crossing treatment selections, when they are warranted by the MUTCD (Ashur and Alhassan 2015). The basic crossing treatment selected in most cases is a marked crosswalk, whose installation mainly depends on speed limit. Additionally, the number of lanes, presence of a raised median or pedestrian refuge island, and vehicle volume are considered before the installation of a crosswalk (Manual on Uniform Traffic Control Devices (MUTCD) 2009). According to a before-and-after study (Mead, Zegeer, and Bushell 2014), both an increase and decrease of the number of pedestrian crashes were observed after the installation of marked crosswalk alone. The authors' study concluded that the impact on pedestrian crashes of a marked crosswalk alone, without additional signs and signals, is inconclusive.

To further improve interactions between drivers and pedestrians at marked midblock crosswalks, signals that alert drivers of the presence of pedestrians within the crosswalk areas have been provided in several locations. The signals, which are predominantly installed in low pedestrian volume areas, can be categorized as: In-pavement Flashing, Circular Flashing Beacons (CFBs), Circular Rapid Flashing Beacons (CRFBs), and Rectangular Rapid Flashing Beacons (RRFBs). At some locations, especially with either high pedestrian volumes and/or high speed limits, Pedestrian Hybrid Beacons (PHBs), or Traffic Control Signals (TCSs) have been installed

(Figure 1) (Bennett, Manal, and Van Houten 2014; Fitzpatrick, Brewer, and Avelar 2014; Van Houten 2011; Pécheux, Bauer, and Mcleod 2009; Prevedouros 2001).



(a) Circular Flashing Beacons (CFB) (b) Circular Rapid Flashing Beacons (CRFB)



(c) Pedestrian Hybrid Beacons (PHB) (d) Rectangular Rapid Flashing Beacons (RRFB)

(e) Traffic Control Signals (TCS)

Figure 1. Typical signals installed at different signalized crosswalks

Most of the signals are activated by using a pushbutton, and each of the signals operates differently when activated. The TCSs operate similarly to the traditional traffic signals at intersections, whereby the lights change from green to yellow, and finally to red before the beginning of the pedestrian crossing phase. The PHBs, on the other hand, consist of units similar to traditional traffic signals, but only flash red when activated. When activated, the CFBs, CRFBs, and RRFBs emit yellow lights in an alternate fashion, which persist for a certain pre-set period. The RRFBs are the most recent technology, whose interim approval was passed in July 2008 and terminated in December 2017 over patent issues. Their interim approval was re-instated in March 2018 (FHWA 2018b, 2018a). Under the interim approval, RRFBs can supplement standard pedestrian crossing warning signs and markings at locations such as pedestrian and school crosswalks where pedestrians' safety is a critical concern (FHWA 2008).

The provision of any crossing location and its associated treatment is expected to improve pedestrian-vehicle interaction by altering behaviors of both pedestrians and drivers. It is through a safety improvement assessment that the effectiveness of the treatment is quantified. Fewer pedestrian crash occurrences or near-crash events after a crossing location treatment has been implemented is the obvious indication of safety improvement. Due to the rarity and randomness of pedestrian crashes and near-crash events, studies have used yielding compliance as a surrogate measure to assess safety improvements (Gates et al. 2016). Yielding compliance is measured by the yielding rate, which is taken as the ratio of yielding drivers to the sum of yielding and non-yielding drivers. Studies have used the yielding rate to compare locations before and after crossing treatments, as well as the variation of the effectiveness of similar treatment types at various locations. Previous studies have revealed a great improvement in pedestrians' safety when flashing beacons supplement marked crosswalks. In fact, when flashing lights are activated, the reported drivers' yielding rates have been tremendously high with varying magnitudes, compared to when the lights are not activated (Al-Kaisy et al. 2016; Fitzpatrick, Potts, et al. 2015; Hunter, Srinivasan, and Martell 2012; Pécheux, Bauer, and Mcleod 2009; Shurbutt and Van Houten 2010).

1.2.Problem Statement

Driver-pedestrian-infrastructure interactions at signalized midblock crosswalks have been extensively studied, whereby different types of statistical models have been developed and conclusions have been made based on the findings. The most evaluated driver behavior at a signalized crosswalks is yielding compliance (Al-Kaisy et al. 2016; Fitzpatrick, Potts, et al. 2015; Hunter, Srinivasan, and Martell 2012; Pécheux, Bauer, and Mcleod 2009; Shurbutt and Van Houten 2010). Yielding is considered as an instance when vehicle(s) have completely stopped or reduced their speeds to allow pedestrian crossing (Hunter, Srinivasan, and Martell 2012). Further, the crossing compliance of pedestrians has been a focus for several researchers interested in pedestrians' behaviors at signalized midblock crosswalks (Brosseau et al. 2013; H. Guo et al. 2014; K. Kim, Made, and Yamashita 2008; Koh and Wong 2014; Rosenbloom 2009; Yanfeng et al. 2010; Zhou et al. 2013).

In modeling drivers' yielding compliance, logistic regressions have consistently been applied (Fitzpatrick, Avelar, et al. 2016; Fitzpatrick, Potts, et al. 2015; Fitzpatrick, Brewer, and Avelar 2014; Kutela and Teng 2018; Porter et al. 2016), since they permit the evaluation of individual crossings rather than aggregated data (Fitzpatrick, Avelar, et al. 2016; Fitzpatrick, Brewer, and Avelar 2014). Among the noted criticisms of the use of logistic regressions, is that in order to apply them, the response variable should be independent of the total number of observed vehicles in a particular crossing (Fitzpatrick, Brewer, et al. 2016). However, this is simply not possible, as the models weigh the crossings proportionally to the number of observed vehicles, while the exposure component among the predictors of proportions is rarely considered. Moreover, the number of yielding drivers is constrained by the number of lanes, while non-yielding drivers are not constrained. Additionally, logistic regression models focus only on whether the vehicles yielded, not the number of vehicles that passed before the yielding occurred. With such a consideration, the Negative Binomial (NB) model has been proposed to model yielding compliance (Fitzpatrick, Brewer, et al. 2016). To apply NB, non-yielding vehicles were counted until a voluntarily yield to pedestrians occurred. Such a data structure resembles a negative binomial experiment.

However, both study methods have several shortcomings. While interrupted traffic flow at signalized crosswalks may exhibit several states and transitions, researchers who have applied logistic regressions assume the existence of two states only: yield and non-yield, with one transition between them (Fitzpatrick, Avelar, et al. 2016; Fitzpatrick, Potts, et al. 2015; Fitzpatrick, Brewer, and Avelar 2014; Porter et al. 2016). One of the shortfalls of this assumption is that it ignores the presence of a "partial-yield" state (Fisher and Garay-Vega 2012; Houten and Malenfant 2008; Zegeer et al. 2005). This state is incorrectly grouped into either of the two above-mentioned states. Secondly, even if only two states exist, the transition between them sometimes does not occur immediately after the pedestrians have pressed the pushbutton (Al-Kaisy et al. 2016; Foster, Monsere, and Carlos 2014; Hunter, Srinivasan, and Martell 2012; Shurbutt et al. 2009). Further, the application of NB by considering vehicle counts faces three major criticisms. First, like logistic regression, this method assumes that traffic flow can exhibit two states only. Secondly, considering the number of vehicles passing by a given location, without describing their arrangement, may yield misleading results. For instance, for a five-lane roadway, the time taken for five vehicles arranged in parallel to pass a location is quite different from the same five vehicles arranged in a series. Lastly, even if vehicle arrangement has no impact, vehicle count is neither an engineering, nor conventional way of delivering technical information. Since at most three states can be observed, multinomial logit can be proposed to model yielding compliance; however, the transition between states is not best presented as choice-based, but rather time-based (Hunter, Srinivasan, and Martell 2012).

In investigating the risks associated with drivers yielding right-of-way at signalized crosswalks, two scenarios, near-miss incidents and partial-yield events, have been reported (Garay-Vega 2008; F. Guo et al. 2010; Hayward 1972; Houten, Malenfant, and Rolider 1985; Matsui, Hitosugi, Doi, et al. 2013; Matsui, Hitosugi, Takahashi, et al. 2013; Sucha, Dostal, and Risser 2017; Voorhees 2017; Zegeer et al. 2005). A near-miss event involves either pedestrians or drivers making an abrupt maneuver to avoid a crash occurrence (Hayward 1972; Houten, Malenfant, and Rolider 1985; Sucha, Dostal, and Risser 2017). Partial-yield incidents involve situations on multilane roadways, in which a vehicle in one lane stops to allow pedestrians to cross, while other vehicles in the adjacent lanes that are driving in the same direction do not stop for the same pedestrians (Houten and Malenfant 2008; Zegeer et al. 2005). Previous studies have successfully associated near-miss incidents and actual crashes but have been unable to associate partial-yield incidents and near-miss events. Therefore, it is still unknown to what extent partial-yield incidents cause risk to crossing pedestrians and vehicle drivers. Moreover, the factors associated with either partial-yield or near-miss incidents at signalized midblock crosswalks are yet to be explored. In addition, since both near-miss, and partial-yield events are very rare, it is not clear whether traditional regression models or rare-events regression models should be applied to model them. Moreover, according to (King and Zeng 2001), traditional regressions tend to result into biased estimates when used to model rare events.

For pedestrian-infrastructure interactions, several studies have been performed to analyze the crossing compliance of pedestrians at signalized crosswalks. In modeling this compliance, two models, multinomial logit and logistic regression, have predominantly been applied (Brosseau et al. 2013; H. Guo et al. 2014; K. Kim, Made, and Yamashita 2008; Koh and Wong 2014; Rosenbloom 2009; Yanfeng et al. 2010; Zhou et al. 2013). When applying these models, studies have considered either temporal or spatial crossing compliance separately, while in reality, spatial and temporal compliance occur jointly (Yanfeng et al. 2010). With that in mind, additional choice based models, such as ordinal models can also be applied. In the previous studies, the choices of model types were not justified, and the performances of the applied models were not assessed.

Several crosswalk features are provided to enable safer interactions between pedestrians and drivers at signalized crosswalks. Most signalized crossing locations are equipped with pushbuttons, which are used to activate flashing lights or request the "walk" phase. A few studies (Carsten, Sherborne, and Rothengatter 1998; Foster, Monsere, and Carlos 2014; Hunter, Srinivasan, and Martell 2012; Levelt 1992) have been conducted to evaluate the frequency of pedestrians pressing pushbuttons before crossing. Indeed, researchers have observed a wide range of variations in pushbutton-pressing compliance; however, most of the studies have not included a wide range of variables, which could explain such variations. Moreover, studies related to supplemental features that provide additional information and directives, such as reminding pedestrians to look for traffic before crossing, as well as directing them to a dedicated location to cross, are scarce. The extent to which pedestrians interact with crosswalk features, such as push buttons, has a role to play in alerting drivers of their presence in crosswalk areas. However, these associated factors for crosswalk activation rates have not been well explored.

Additionally, the influence of other information provided to pedestrians and drivers at crossing locations has not been a main topic of interest. For instance, the influence that pavement markings, signs, and other features have on pedestrians' decision making before and during crossing is not well researched. Apart from pedestrian concerns, drivers are supposed to yield right-of-way to pedestrians at certain predefined distances marked by yield lines. According to the MUTDC (Manual on Uniform Traffic Control Devices (MUTCD) 2009), under normal traffic and land use conditions, yield lines should be positioned not less than 40ft or greater than 180ft from crosswalk signals. Furthermore, under special scenarios, yield lines can be as close to the signal as possible. Although studies have attempted to evaluate the yielding compliance of drivers at varying distances from yield lines to crosswalks, two shortfalls are observed. First, studies have focused on the presence of a yield line only (Houten, Malenfant, and McCusker 2001; Van Houten et al. 2002; Samuel et al. 2013); thus, neglecting the combined effects that other features might contribute. Second, as a result of focusing on the presence of a yield line only, descriptive analysis has been the dominant methodology used in these studies. Further, the influence of various distances from the marked strips to the yield lines has not been explored.

1.3.Research Objectives

This research seeks to analyze driver-pedestrian-infrastructure interactions at signalized midblock crosswalks. In so doing, descriptive analyses and statistical tests, as well as models that evaluate crosswalk users' interactions are developed. The following are specific objectives of the study:

- The introduction and implementation of an improved methodology for evaluating the yielding compliance of drivers at signalized midblock crosswalks. In this approach, the durations of the transitional states involved while drivers are yielding right-of-way to pedestrians are quantified and modeled. These transitional states are defined as non-yield, "partial-yield," and full yield. Multistate models, which are a family of hazard-based models, are introduced to associate yielding compliance and other co-variates.
- The evaluation of the risks associated with driver-pedestrian interactions at signalized midblock crosswalks. Partial-yield and near-miss events are the main risks given special

attention. The methodology used for analysis, relationship between the two events, and associated factors for partial-yield and near-miss incidents at signalized midblock crosswalks are presented.

- The analysis of pedestrian-infrastructure interactions, whereby the spatial and temporal crossing compliance of pedestrians, effective use of crosswalk features, and influence of crosswalk features on pedestrians' behaviors are the main focus. In this objective, the study presents an assessment of the methodological alternatives for the analysis of the spatio-temporal crossing compliance of pedestrians at signalized midblock crosswalks, whereby several models are proposed and evaluated based on several performance measures. Moreover, the influential factors for pushbutton utilization are evaluated.
- The evaluation of the influence of crosswalk features, traffic conditions, and pedestrians' characteristics on drivers' spatial yielding compliance at signalized midblock crosswalks. The models that predict the spatial yielding compliance are presented, whereby the combined influences of crosswalk features are compared to the combined influences of non-crosswalk features against that of non-crosswalk features.

1.4.Research Scope

This dissertation considers yielding as both a voluntary action by drivers, and as stated by law. According to Nevada (and most states') laws on driver yielding, a driver is required to yield to pedestrians who are already in a crosswalk. The NRS 484B.283 (a) states that "*When official traffic-control devices are not in place or not in operation, the driver of a vehicle shall yield the right-of-way, slowing down or stopping if need be so to yield, to a pedestrian crossing the highway within a crosswalk when the pedestrian is upon the half of the highway upon which the vehicle is traveling, or when the pedestrian is approaching so closely from the opposite half of the highway* *as to be in danger"* (NRS: CHAPTER 484B - RULES OF THE ROAD n.d.). Thus, regardless of whether flashing lights are activated, or pedestrians are on the curb waiting to cross, drivers are not by law required to yield. However, in this study, drivers are expected to yield the right-of-way when pedestrians are either waiting to cross or already crossing, regardless of flashing light status. Per this study, yielding is defined as the speed reduction or stoppage of all vehicles so that pedestrians may use the crosswalk. This study also considers both pedestrians' and drivers' behaviors in their utmost possible natural behaviors.

Geographically, this study is based in Las Vegas, Nevada. The state of Nevada, with 2.76 pedestrian fatalities per 100,000 people, has been ranked sixth nationwide. Moreover, pedestrian fatalities within the state, particularly in the Las Vegas Metropolitan area, are escalating each year. The most recent statistics indicate that between 2010 and 2016, a total of 441 fatal and more than 700 injury-causing pedestrian crashes have occurred on the Nevada roadway network (NDOT n.d.; Strategic Highway Safety Plan 2016). The data for this study are collected from twenty (20) signalized midblock crosswalks located in Las Vegas, Nevada. The crossing locations that are considered are equipped with either Circular Flashing Beacons (CFBs), Circular Rapid Flashing Beacons (CRFBs), Rectangular Rapid Flashing Beacons (RRFBs), Pedestrian Hybrid Beacons (PHBs), or Traffic Control Signals (TCSs) installed at locations that are specifically dedicated for pedestrians to cross.

1.5.Study Approach

To attain the stated objectives, this study is divided into six main parts, which are introduction, literature review, methodology, descriptive analysis, model development and results discussion, and conclusion and recommendations (Figure 2). Each part of the study has its own specific purpose. In the introduction portion, the study motivation, problem statement, objectives, study scope, and contributions of the findings are presented. The literature review establishes the gap in the literature not covered by previous studies, by considering the methodologies and data used.

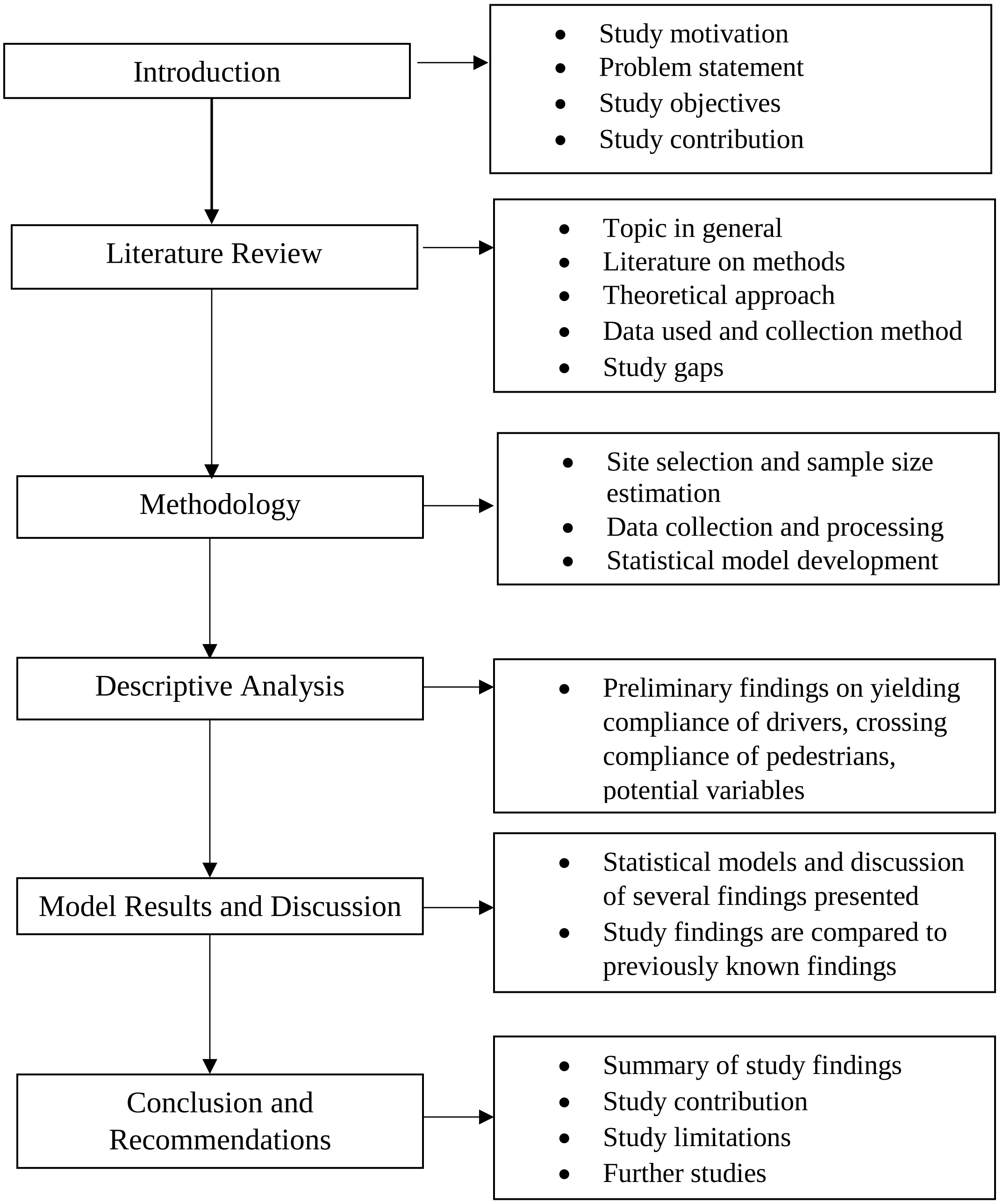


Figure 2. Methodological framework

The methodology section avails the approach proposed in this study. In this part, the site selection, sample size estimation, data collection and processing procedures, identification of potential variables, and statistical models are developed. Further, a descriptive analysis of potential variables is performed to provide the preliminary findings. The descriptive analysis results may not be the final results, since they are rarely generalized; thus, the statistical models are performed for general inference. Through the model results and discussion section, several findings are discussed, which facilitate the conclusion and recommendations section. In the conclusion and recommendations section, study limitations are also presented.

1.6.Study Contribution and Application of Findings

An important aspect of the study is the application of its findings in improving or updating existing conditions. This part of the dissertation presents the expected study contributions and applications of the findings. The contributions and applications of this study are divided in terms of academics, practitioners, and the public.

1.6.1. Academic contributions and application

- This study introduces to the body of literature the hazard-based multistate model for modeling the yielding compliance of drivers. Contrary to the previously developed models designed to study the yielding compliance of drivers, the multistate model developed in this study considers transitional states and their corresponding transition durations while drivers are yielding right-of-way to pedestrians. Therefore, this study presents to researchers a model that not only evaluates the probability of drivers to yield, but also the transitional states involved, and the time taken to yield. The developed model enables researchers to evaluate the associated factors for drivers' yielding right-of-way to pedestrians in a more realistic way.

The developed multistate models can be extended and used to model other traffic incidents that have progression sequences. A typical example, in which such multistate models can be applied, is in the modeling of highway incident clearance duration, where a sequence of operations that moves from incident occurrence, to incident detection, and incident clearance is expected. Such a process can be modeled by a progressive multistate model, which is in the family of hazard-based models, similar to permanent illness-death models, applied in this study.

- The study also presents the framework to assess the modeling methodologies for the spatiotemporal crossing compliance of pedestrians at signalized midblock crosswalks. In this case, the study develops and compares three models: multinomial logit, ordered logit, and logistic regression. Several performance measures, including the models' prediction accuracies and information-based criteria are applied to determine the best modeling methodology. Based on these performance measures, the best modeling methodology is proposed. This approach is applied since both spatial and temporal crossing compliance occur jointly. In the transportation engineering field, and more specifically traffic safety, there are scenarios that occur jointly, but are modeled separately. A typical example is the modeling of crash injury severities, where in the same crash there might be individuals who die, and others who are injured, but the crash is generally categorized as fatal. The approach in this study can be a revelation for modeling such types of scenarios.
- The near-miss and partial-yield incidents are rarely observed at signalized midblock crosswalks, as shown in this study. Due to such rarity, the modeling of these incidents should be carefully handled to avoid reporting biased results. This study proposes a methodology for modeling these rare events, compared to the traditional methods, and proposes the best

model based on several applied performance measures. Modeling similar rare events in transportation engineering has not been extensively considered. This study, therefore, paves the way for researchers in transportation engineering to have a proper approach/consideration when dealing with rare incidents, such as secondary crashes, mass casualty traffic incidents, etc.

1.6.2. Contributions and application to practitioners

- Engineers, city planners, and researchers may use the developed hazard-based yielding compliance models to evaluate drivers' yielding compliance at pedestrian crossing locations, for different crosswalk treatments at the same location, or for before-and-after studies for the same location, when there is a change in a pedestrian crossing treatment. The developed multistate models may be used by traffic engineers, planners, and researchers to evaluate changes in the spatiotemporal states of traffic flow, given any change in crossing treatment performed at a location.
- This study evaluates the influence of various crosswalk features, such as signs and markings, traffic related variables, and pedestrian related factors on the safe utilization of signalized midblock crosswalks. These include factors influencing yielding compliance, as well as crossing compliance. Engineers, planners, and policy makers can use the findings from this study when establishing crossing locations for pedestrians. For instance, although flashing lights and pedestrian signals have been provided at most pedestrian crossing locations across the United States, a major concern is related to how to enable the pedestrians to always activate the lights before crossing, and to cross at dedicated locations. Understanding the determining factors associated with pushbutton activation is a great step towards solving this problem. Determining the influence of signs and markings positioned at pedestrian crossings

not only provides the opportunity for installing similar signs in other locations with similar characteristics, but also the incorporation of additional messages to educate pedestrians.

1.6.3. Contributions and Application to the general community

- Through this study's findings and recommendations, the community can benefit from improved pedestrian safety at crosswalks. For instance, a better design of the crosswalk features that facilitate safe interactions between pedestrians and drivers can be proposed. Furthermore, modifications of crosswalk features that have no influence on safe interactions between crosswalk users and drivers may be proposed.

1.7.Dissertation Organization

The dissertation is organized into seven chapters, whereby this introduction chapter is followed by a literature review, in which crosswalk signalization is introduced; pedestrian-infrastructure interaction is deeply discussed; pedestrian-driver interaction is extensively reviewed; and finally a review of the statistical models used in previous studies is presented. Chapter three presents, in detail, the methodology applied in this study, whereby the experiment design is presented, and hazard-based models and binary choice models are discussed in terms of concept, estimations, and interpretations. Chapter four of the report presents the descriptive analysis of the collected data, followed by model results; their discussion is in chapter five. Chapter six finalizes the main body of the report by presenting the conclusions and recommendations, as well as future works that were not covered in this study. Lastly, references are presented.

CHAPTER 2: LITERATURE REVIEW

This chapter covers the literature review, starting from midblock crosswalk signalization, pedestrian-infrastructure interaction, pedestrian-drivers interaction, to statistical models utilized in the previous studies. Through this chapter, various studies that explored the aforementioned interactions are reviewed and the strengths and weakness of their findings are analyzed in order to identify the gap in the existing literature.

2.1.Midblock crosswalks signalization

According to the manual on uniform traffic control devices (MUTCD) (Manual on Uniform Traffic Control Devices (MUTCD) 2009), a signalized midblock crosswalk is any signalized crosswalk that is between two signalized intersections. These crosswalks can be signalized by using several signal types depending on existing factors, such as traffic volume, pedestrian volume, and land use, to mention a few. At least five types of crosswalk signalizations are available. These are, Circular Flashing Beacons (CFBs), Circular Rapid Flashing Beacons (CRFBs), Rectangular Rapid Flashing Beacons (RRFBs), Pedestrian Hybrid Beacons (PHBs), and Traffic Control Signals (TCSs) (Bennett, Manal, and Van Houten 2014; Fitzpatrick, Brewer, and Avelar 2014; Van Houten 2011; Pécheux, Bauer, and Mcleod 2009; Prevedouros 2001). The operation of each of the signal is different from the other. This section provides a detailed explanation of the operations of each of the signal types.

2.1.1. Circular Flashing Beacons (CFBs)

The Circular Flashing Beacons (CFBs) consist of alternating flashing lights housed in a circular tunnel visor, similar to those used on traffic lights at signalized intersections. When activated, CFBs flash in an alternate fashion at a predefined rate and time. After that time, the entire system becomes dark until the next activation. Figure 3 shows the CFBs installed at Maryland Pkwy and Reno Avenue in Las Vegas, Nevada.

Figure 3. Circular Flashing Beacons (CFBs)

2.1.2. Rectangular and Circular Rapid Flashing Beacons (RRFBs) (CRFBs)

The RRFBs are formed by a rectangular bar, while the CRFBs are incased in circular tunnel visors similar to those used on traffic lights at signalized intersections. Both RRFBs and CRFBs use light emitting diodes (LEDs) that flash in a similar fashion as those on emergency vehicles (Van Houten 2011). The flashes last for a predefined time, usually 30 to 40 seconds. Figure 4 shows CRFBs (a) and RRFBs (b) installed on two roadways in Las Vegas, Nevada. The CRFB is on Valley View Blvd near El Conlon Ave, while the RRFB is on Flamingo Road near Cameron Street.

(a). Circular Rapid Flashing Beacons (CRFBs) (b). Rectangular Rapid Flashing Beacons (RRFBs)

Figure 4. Typical CRFBs and RRFBs

2.1.3. Traffic Control Signals (TCSs)

Traffic Control Signals (TCSs) are similar to those found at most of signalized intersections. They consist of green, yellow, and red lights. When there is not a pedestrian a crossing phase, the TCSs display green lights, which enables the vehicular traffic phase. Upon activation by a pushbutton, the green phase changes to red after certain preset time, and the pedestrian crossing phase is activated. During this phase, a pedestrian sign is displayed, followed by either a countdown or a hand sign on the same box displaying a pedestrian walking. A typical TCS installed on Maryland Pkwy near Del Mar Street is presented in Figure 5 below.

Figure 5. Traffic Control Signals (TCSs)

2.1.4. Pedestrian Hybrid Beacons (PHBs)

The Pedestrian Hybrid Beacons (PHBs) consist of a unit similar to a traditional traffic signal, but with a red–red instead of yellow (amber) format. When not activated, the unit remains dark, but when activated, the unit flashes yellow to alert drivers to prepare to stop. The unit then changes to solid red, indicating to drivers that they should stop, and at this moment the ‘WALK’ symbol for pedestrians is illuminated. After a certain preset time, the unit starts flashing red, which indicates that drivers should come to a complete stop and look for crossing pedestrians; if there are no pedestrians, drivers should proceed. The red flashing continues for a certain preset time then turns dark. Figure 6 shows the progression of lights when PHBs are activated, while Figure 7 shows a typical PHB installed on Sahara Avenue near 15th Street in Las Vegas, Nevada.

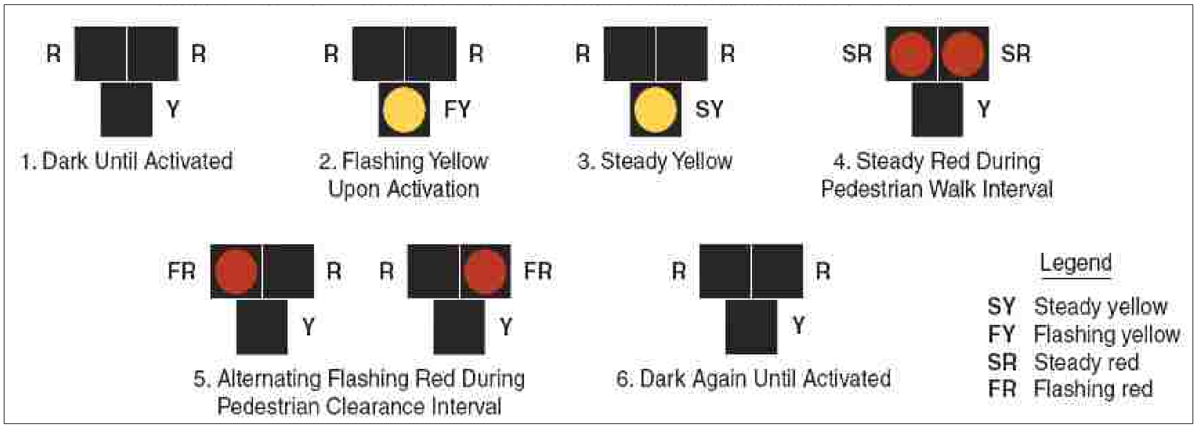


Figure 6. Pedestrian Hybrid Beacons (PHBs) operation

(Manual on Uniform Traffic Control Devices (MUTCD) 2009)

Figure 7. Installed Pedestrian Hybrid Beacons (PHBs)

2.2.Pedestrian-infrastructure interactions

Before pedestrians have interacted with drivers/vehicles, they first interact with crosswalk features. Although this interaction may well define the interaction between pedestrians and drivers, it has not been extensively studied, especially for midblock crosswalk settings. There are a few studies that have focused on signalized intersection settings. Signalized midblock crosswalks and signalized intersections have several similarities, which include dedicated pedestrian crossing zones and pedestrian crossing phases (for PHBs and TCSs), as well as the presence of pushbuttons for activating the lights to request crossing and the presence of median refuges, to mention a few. However, the operations of various crosswalks are significantly different. For instance, for signalized intersections, whether a pedestrian has pushed a button or not, there is a predefined pedestrian crossing phase for each approach. The situation is very different for signalized midblock crosswalks, in the sense that the pedestrian is supposed to request a crossing phase. This is predominantly done by using a pushbutton, although other means such as automatic detectors are in use in some locations (Hughes et al. 2000; Nambisan et al. 2009). Several assessment criteria may define proper pedestrian-infrastructure interactions. Included in the related literature are pushbutton utilization, inappropriate crossing, pedestrian delay, and looking before crossing.

2.2.1. Pushbutton utilization

Pushbuttons remain the traditional way of activating warning lights. Pushbuttons are either mounted on a pole with the traffic signals/flashing lights or placed on a different pole close to the crosswalk if the pole with the traffic signals/flashing lights is located at a distance from the dedicated crossing location (marked stripes), Figure 8. Pushbuttons of different designs are purposely provided at signalized crosswalks so that pedestrians may push to request a crossing phase. However, studies have revealed that pedestrians frequently do not push the buttons (Carsten, Sherborne, and Rothengatter 1998).

Figure 8. Types of pushbuttons, message, and placements

A study which involved three countries, France, the United Kingdom, and the Netherlands (Levelt 1992), was among the early studies that focused on the pedestrian activation of pushbuttons at signalized intersections. This study applied a survey questionnaire, as well as an observational survey by using video cameras, to evaluate pedestrians' behaviors towards the use of pushbuttons. Through survey questionnaire results, it was revealed that in the United Kingdom, 40% to 50% of the respondents said they always press the pushbutton, while in the Netherlands, only 34% of respondents (68 out of 201) provided a similar response. A large percentage, 41% of respondents, in the Netherland said they would push the button provided that no one else had done so. The percentage of respondents who said they never pressed the pushbutton varied from 12% in the

Netherlands to between 11% and 22% in the United Kingdom. However, the data processed from video revealed a very different trend; for instance, in the United Kingdom, the observed button pressing rate ranged between 14% and 35%, while in France 18% of the crosswalk users pressed the pushbutton.

In the United States, several studies have been devoted to this topic; those worth mentioning include studies in Bend, Oregon (Ross, Serpico, and Lewis 2011), in Portland, Oregon (Foster, Monsere, and Carlos 2014), in Saint Petersburg, Florida (Hunter, Srinivasan, and Martell 2012), in Santa Monica, California (Morrissey 2013), in Montana (Al-Kaisy et al. 2016), and in Virginia (Dougald 2015). These studies reported varying levels of pushbutton activation for different crosswalk signals.

In Bend, Oregon, a before-and-after study (Ross, Serpico, and Lewis 2011) assessed driver yielding rates at three crosswalks on Bend Parkway (Reed Lane and Badger Road) and Greenwood Avenue at NE 12th Street. In the after period, all crosswalks were equipped with RRFBs. Bend Parkway, which is a four-lane roadway with a center median, with bike lanes and sidewalks, has a speed limit of 45mph. Further, the posted speed limit for Greenwood Avenue, which is a five-lane roadway with a two-way center left turn lane, was 35mph. In addition to the pushbutton, an audible device was provided, so that when a pedestrian pressed the button, the following message was heard: “Lights are on to cross the Parkway. Traffic may not stop.” The authors used both video cameras and printed sheets to record pedestrian-driver interactions. The video recording lasted for at least two days, whereby 78 crossings incidents were recorded at Reed Lane, 60 at Badger Road, and 51 at NE 12th Street. At two locations, Badger Road and NE 12th Street, staged pedestrians, who were instructed to press pushbutton all the time before utilizing the crossing, were used. Meanwhile, at Reed Lane with 78 general pedestrians, whereby 64% (50 out of 78) of them were bicyclists, 50% (25 out of 50) did not activate the flashing lights. At the same location, 75% of the remaining crosswalk users (21 out of 28 pedestrians) pressed the button to activate the flashing lights. Although the authors provided their analysis on the utilization of pushbuttons, this study had a small sample size, which makes it difficult to draw a concrete conclusion. In addition, the authors did not associate the compliance of crossers' effective utilization of the pushbutton with any other variable, apart from showing the difference in compliance between pedestrians and bicyclists.

Another study (Hunter, Srinivasan, and Martell 2012) in Saint Petersburg, Florida, evaluated the performance of RRFBs at a trail where most of the users were bicyclists. The trail crosses a minor arterial street (22nd Avenue North) that has two lanes in each direction, and was estimated to have 15,000vpd, with a posted speed limit of 40mph. The trail serves approximately one to two thousand users per day, almost 80% of whom were bicyclists. This study was a before-and-after study, whereby during before period, no crosswalk treatment existed. On the other hand, during the after period, a marked crosswalk, equipped with RRFBs activated by pushbutton, was put in place. The analysis of 400 trail users, who were recorded using an elevated camera during the after period, revealed that only 32% activated the flashing lights using the pushbutton. Among the 68% of trail users who did not activate lights, only 19% arrived while a previous user had already activated the lights; this indicates that 49% of the users did not activate the lights. Although the statistics were not provided, the authors observed that, compared to pedestrians, bicyclists were less likely to activate flashing lights. Bicyclists tended to wait for an available gap by approaching the crosswalk with reduced speed. With that low rate of flashing light activation, the city installed the reminder "PUSH BUTTON TO ACTIVATE BEACONS" on the stop sign's pole. However, the reminder did not solve the problem, as the activation rate continued to be relatively the same.

Although the study analyzed the extent to which trail users pushed the button to activate lights, it did not statistically associate the pushbutton activation tendency with any other explanatory variables.

A relatively higher activation rate was observed in Portland, Oregon (Foster, Monsere, and Carlos 2014). Their study was aimed at evaluating driver and pedestrian behaviors at Danish-offset midblock crosswalks equipped with RRFBs located on two multi-lane roadways: Barbur Boulevard and Beaverton-Hillsdale Highway, with 35 and 40 mph speed limits, respectively. Being Danish-offset (Z-crossing), the crosswalks had two stages, in such a way that pedestrians could cross one stage (direction of traffic flow) and face the incoming traffic, before crossing the second stage. Sixty-two hours of video recording was performed at the two sites, where a total of 351 pedestrian crossing incidences were recorded. This study reported that out of 196 pedestrian crossing incidences, 173 (92%) of the pedestrian's crossing pressed the pushbutton to activate the flashing lights at the Southwest Barbur Boulevard site, while 83% (123 out of 155) did the same at the Beaverton-Hillsdale Highway site. Further analysis revealed that the activation rates across the two sites were higher, 94% (160 of 170) and 89% (112 of 126), respectively, when there were some incoming vehicles, compared to 72% (13 of 18) and 48% (11 of 23) when there were no incoming vehicles. The authors performed a two-sample z-test of proportion and found that the activation rate was statistically significantly different in the presence and absence of incoming vehicles. Furthermore, among the incidences where no activations were performed, about 5% (8 out of 173) and 4% (6 out of 155) at Barbur Boulevard and Beaverton-Hillsdale Highway, respectively, of pedestrians arrived at the crosswalks while the RRFBs were still flashing from a previous actuation. However, the authors agreed that this sample was too small to draw any tangible conclusion. The authors hypothesized that the activation rate may have depended on the speed limit and the crossing length, as the roadway with the higher speed limit and longer crossing length was observed to have a higher activation rate. Nonetheless, no statistical association of the hypothesized factors was developed.

Additional research related to pushbutton usage presented varying results. In Santa Monica, California (Morrissey 2013) found varying activation rates when RRFBs and CRFBs were used. At one site, the activation rates for CRFBs were 92% higher compared to RRFBs at 85%; at another location, RRFBs with an 80% activation rate outperformed CRFBs with a 63% activation rate. A two-site study in the state of Montana (Al-Kaisy et al. 2016), which was performed at King Avenue and Kagy Boulevard, found that the activation rates at the crosswalks equipped with RRFBs were about 57% and 81%, respectively. In another study (Brewer, Fitzpatrick, and Avelar 2015), the maximum flashing light activation rate was 94%, though at some locations a low number of pedestrians (e.g., six) was observed; thus, the statistics, in terms of percentage, might be misleading. Other research conducted in Virginia showed that the percentage of trail users who activated flashing lights was observed to be 23.8% after three weeks, 29.3% after five months, and 27.3% after a year (Dougald 2015). A further study reported that an elevated speed limit resulted into a high pushbutton pressing rate (Fitzpatrick, Avelar, et al. 2016). Their study found that 91% of the pedestrians who crossed at crosswalks with PHBs pushed the pushbutton, with the rates for 45mph roadways outpacing 40mph or less roadways.

In than attempt to motivate pedestrians to use a pushbutton, researchers adopted an illuminated pushbutton. A before-and-after study (Huang and Zeegar 2001) using four signalized intersections in Windsor, Ontario, found that even after installing illuminated pushbuttons, there was no statistically significant increase of pedestrians pushing the buttons. As a matter of fact, the pushbutton pressing percent declined, from 16.9 % to 12.7%, after the installation of the illuminated pushbuttons. The same study gave possible reasons for this, among which were the pedestrians arriving when there was a walk signal, and the pedestrians utilizing the available gap in opposing traffic, even if the parallel traffic had the red light.

As a result of unsatisfactory pushbutton activation rates in Montana, a study (Al-Kaisy et al. 2016) recommended positioning them at locations that are more practically possible for access; however, other researchers introduced automatic pedestrian detection devices. In fact, two studies (Hughes et al. 2000; Nambisan et al. 2009) evaluated automated pedestrian detection systems that trigger crossing phases. Although there were fewer pedestrian-vehicle conflicts, the system faced a high number of false calls, in which the lights were triggered when a pedestrian was not intending to cross, as well as missed calls, in which the crossing pedestrian was not detected (Hughes et al. 2000). The conclusion of this study calls for a thorough pedestrian activity study on the crossing patterns and proportion of through to crossing pedestrians, among other factors, prior to deploying automatic pedestrian detection.

2.2.2. Inappropriately crossing/jaywalking

All the marked crosswalks have locations that are dedicated for pedestrians to cross through. They are typically marked by either several lines parallel to the vehicular traffic flow (Figure 9) or two lines perpendicular to the vehicular traffic flow. In addition, pedestrians can use the area between vehicle yield lines.

Figure 9. Typical crosswalk markings

Inapropriate roadway crossing may be described in different ways, including jaywalking and crossing outside the dedicated location when vehicles have stopped. Jaywalking is a terminology describing inappropriate crossing of the roadway 10 ft outside of a marked or unmarked crosswalk at either an intersections or midblock (Sisiopiku and Akin 2003; Zheng et al. 2015), where one does not consider the state of the incoming traffic. However, this definition does not describe whether a marked crosswalk includes the yield line. Although crossing within a marked or unmarked croswalk is considered a permissible crossing by definition, there are some instances in which innappropriate crossing may happen. For instance, crossing within a marked crosswalk while it is not a pedestrian phase. A study by (Foster, Monsere, and Carlos 2014) is among those focused on the way pedestrians effectively used dedicated crossing locations. The study found that 70% (155 of 221) properly used crosswalks by crossing within the marked stripes. Moreover, about 15% (33 out of 221) of pedestrians crossed within the legal limits, but not within within the marked stripes, while the rest (15%) jaywalked.

A more rigorous and intensive study on pedestrian crossing compliance was performed in East Lansing, Michigan (Sisiopiku and Akin 2003). Their study used video and survey questionnaires to evaluate pedestrian movements within different crossing facilities, including signalized intesections, and signalized and non-signalized midblock crosswalks. The crossing compliance rate (CCR) was a measure of compliance, which further categorized into temporal crossing compliance rate (TCCR) and spatial crossing compliance rate (SCCR). The SCCR was defined as the ratio of crossing incidences that occurred within dedicated crossing locations (marked stripes, plus 3 ft on either side), to total crossing incidences within a crosswalk influence area. TCCR was defined as the ratio of crossing incidences in which pedestrians waited for their crossing phase, to total crossing incidents. Through analysis of observational survey data, it was found that signalized intersections had the highest SCCR 82.8%, followed by midblock crosswalks 71.2%, and unsignalized intesections 67.5%, while non-striped midblock crosswalks had the least SCCR at 64.2%. The survey questionnaire, which was responded to by 711 respondents, results revealed that only 5.8% of the respondents never crossed at non-designated locations, while 4.2% always crossed in non-designated locations. Among the reasons for crossing within non-designated locations, convenience (39.5%), light traffic (28.7%), and saving time (25.9%) were the top three reasons.

Another study (Nambisan et al. 2009) revealed that, as a result of an installed signalized midblock crosswalk with an automatic pedestrian detector, the number of diverted pedestrians increased from 0 to 14. This study, however, neither explained the specific crossing zone of the diverted pedestrians, nor whether the diverted pedestrians were considered jaywalkers before the installation of the signalized midblock crosswalk. Compared to the midblock crosswalks, intersections are more prone to jaywalking. A study around the campus of the University of

Florida (Zheng et al. 2015) found that, on average, there were about one to four jaywalkers per minute at five crosswalks within the campus. The study further revealed a positive correlation between jaywalking and the presence of a bus stop, the distance between crosswalks, and traffic volume, while a negative correlation was observed for high traffic volume and longer crossing distance.

2.2.3. Signs at the crosswalks

Looking for incoming traffic before starting to cross can be translated as a safe crossing behavior. For Danish-offset (Z-shaped) crosswalks, pedestrians are forced to face the incoming vehicles before crossing the second stage of the roadway; however, there is no mandatory mechanism that forces pedestrians to do the same before crossing.

Figure 10. Look before crossing and use crosswalk sign

In recent years, most of the signalized midblock crosswalks have had signs that display "LOOK BEFORE CROSSING" (Figure 10), which remind pedestrians to look for incoming vehicles before crossing. It is not clear whether the presence of the signs adds anything to pedestrians' behaviors of looking before crossing. Pedestrians' compliance on looking before crossing has been reported in previous studies for midblock crosswalks and intersections. For instance, a study by (Nambisan et al. 2009) reported that all observed pedestrians (84) looked for the incoming traffic before crossing the first and second half of the roadway. Their study was performed using data collected at a midblock crosswalk equipped with an RRFB. However, (Nambisan et al. 2009) did not describe whether there were any signs instructing pedestrians to look before crossing. In an different study, 6.4% of pedestrians who crossed at a signalized intersection did not look for incoming vehicles (Hamidun et al. 2016). Their analysis revealed that there was no statistically significant difference across gender. The reviewed studies did not reveal the influence of signs on pedestrian compliance, either to look before crossing or to use the crosswalk. None of the previous studies evaluated the influence of the presence or absence of the directive signs at the crosswalks. It is not clear whether pedestrians looked for incoming traffic in a natural manner or were influenced by the presence of the signs.

2.3. Pedestrians-Drivers interactions

2.3.1. Yielding compliance

Among intensively analyzed areas of pedestrian-driver interaction is yielding compliance, since it has been extensively used as a measure of effectiveness of signalized crosswalk treatments. Yielding compliance has been used to compare the drivers' behaviors before and after installations of signalized crosswalks, as well as for different types of crosswalk signalizations. Some studies define yielding compliance per lane, while others define it per pedestrian crossing incidence.

Yielding compliance is measured by the yielding rate, which is defined as the ratio of number of vehicles yielded to the total number of vehicles observed (yielded and non-yielding vehicles) (Brewer, Fitzpatrick, and Avelar 2015; Fitzpatrick, Kay, et al. 2016; Fitzpatrick, Brewer, and Avelar 2014; Foster, Monsere, and Carlos 2014; Turner et al. 2006). The definition of number of yielding vehicles, however, differs from study to study. For instance, if vehicles were in a platoon (Brewer, Fitzpatrick, and Avelar 2015; Fitzpatrick, Brewer, and Avelar 2014) considered the number of yielding vehicles as only the vehicles that were in the front row, since the drivers who yielded behind had no opportunity to decide on whether or not to yield to the pedestrian. Therefore, under this approach, the maximum number of yielding vehicles was equal to the number of lanes available.

A study by (Foster, Monsere, and Carlos 2014), on the other hand ,recorded the observation of each individual vehicle as it yielded or did not to a pedestrian, as required by Oregon law. This study did not elaborate on the vehicles yielding behind other yielded vehicles. Although (Porter et al. 2016; Shurbutt et al. 2009) did not explicitly define yielding rate, their study collected the number of drivers who yielded and those who did not yield to pedestrians. In accordance to (Van Houten, Ellis, and Marmolejo 2008), the percentage of yielding and non-yielding vehicles were scored. Similar to (Foster, Monsere, and Carlos 2014), a study by (Van Houten, Ellis, and Marmolejo 2008) scored a vehicle as non-yielding if it passed in front of a crossing pedestrian when it was able to stop. However, according to (Van Houten, Ellis, and Marmolejo 2008), for a vehicle to be recorded as non-yielding, a pedestrian must have placed at least one foot in the crosswalk, which is in accordance to Florida law. Their study used a staged pedestrian for most cases, whereby, a pedestrian would step into the travel lane and see whether a driver would stop; if the driver stopped, a pedestrian would move to the next lane. In case an un-staged pedestrian used the crosswalk, a similar approach was used to score the yielding and non-yielding vehicles. With that procedure being followed, this approach could better explain the yielding compliance by travel lane. However, a researcher could not evaluate the yielding compliance of the middle lane as a starting lane to cross, since all of the crossings started either at the outer or inner lanes. On the other hand, the yielding compliance by (Potts et al. 2015) was in terms of pedestrian crossing incidences, in a sense that vehicles were considered yielded, when all vehicles in one direction of travel stopped or reduced speeds for pedestrians to cross.

A wide range of driver yielding compliance for different signalization types, when they are either active or inactive, has been reported by studies performed in cities across the United States and Canada. For studies using RRFB, (Shurbutt and Van Houten 2010) looked at 22 sites, most of which were located in St. Petersburg, Florida, and reported 72% to 96% of driver yielding compliance, while (Pécheux, Bauer, and Mcleod 2009) reported a maximum of 60% to 70% for day and night, respectively, using two sites located in Miami, Florida. With most of the 22 sites included in their study located in Garland, Texas, a range of 37% to 89% of driver yielding compliance was reported by (Fitzpatrick et al. 2014; Fitzpatrick, Brewer, and Avelar 2014). Additionally, studies by (Bennett, Manal, and Van Houten 2014; Domarad, Grisak, and Bolger 2013) reported 69% and 98% of driver yielding compliance in their studies in Michigan, USA and Alberta, Canada, respectively. For CRFBs, the reported yielding compliance varied from 63% to 92% when activated, and 57% to 83% when not activated during the day time. On the other hand, during the night time, the range varied between 65% to 90% when activated and 35% to80% when not activated for two sites in Santa Monica, California (Morrissey 2013). In Arizona, Texas, and Wisconsin, (Fitzpatrick, Avelar, et al. 2015) reported an average of 67% yielding compliance during the day time and a relatively higher percentage (69%) during the night time.

Installation of PHBs has been shown to result into very high yielding compliance. When considering staged pedestrians only, the PHB resulted into 90% and 99% yielding compliance for near side and far side directions of traffic flow, respectively. The near side was the side where pedestrian was originating (Brewer, Fitzpatrick, and Avelar 2015).On the other hand, when all pedestrians were considered, the yielding compliance fell to 76% and 52%. A range of 93% to 99% yielding compliance, which is roughly equivalent to that of TCS, was reported by (Fitzpatrick et al. 2006). Another study in four cities, Austin, Houston, San Antonio, and Waco, Texas, found that the yielding compliance of PHBs varies between 72% and 94% (Fitzpatrick, Brewer, and Avelar 2014). As was expected, since TCS are similar to the traffic signals located at most signalized intersections, TCS were found to have the highest overall yielding rate (98%) in Texas, compared to other treatments. This study constituted seven sites located in of four cities: one site in Austin, four sites in Dallas, and two sites Houston, whereby 100% yielding rate was observed in Austin, 99% in Dallas, and 95% in Houston.

Studies have further evaluated not only whether drivers yielded to pedestrians, but also the way they yielded. A study by (Porter et al. 2016) divided yielding into two types: soft and hard yielding; defining hard yielding as the one involving vehicles that stopped abruptly. Their study, however, found that a very small percentage (0.8%) of drivers that were categorized as “hard-yield” while “soft-yield” had a total of 76.3%. Their study ignored the “hard-yield” drivers and continued with only two options of yield and non-yield for further analysis.

2.3.2. Risks associated with drivers yielding right-of-way to pedestrians

In investigating risks associated with drivers yielding right-of-way at signalized crosswalks, two scenarios, near-miss incidents and partial-yield events, have been reported (Garay-Vega 2008; F. Guo et al. 2010; Hayward 1972; Houten, Malenfant, and Rolider 1985; Matsui, Hitosugi, Doi, et al. 2013; Matsui, Hitosugi, Takahashi, et al. 2013; Sucha, Dostal, and Risser 2017; Voorhees 2017; Zegeer et al. 2005). A near miss event involves either pedestrians or drivers making an abrupt maneuver to avoid crash occurrence (Hayward 1972; Houten, Malenfant, and Rolider 1985; Sucha, Dostal, and Risser 2017).

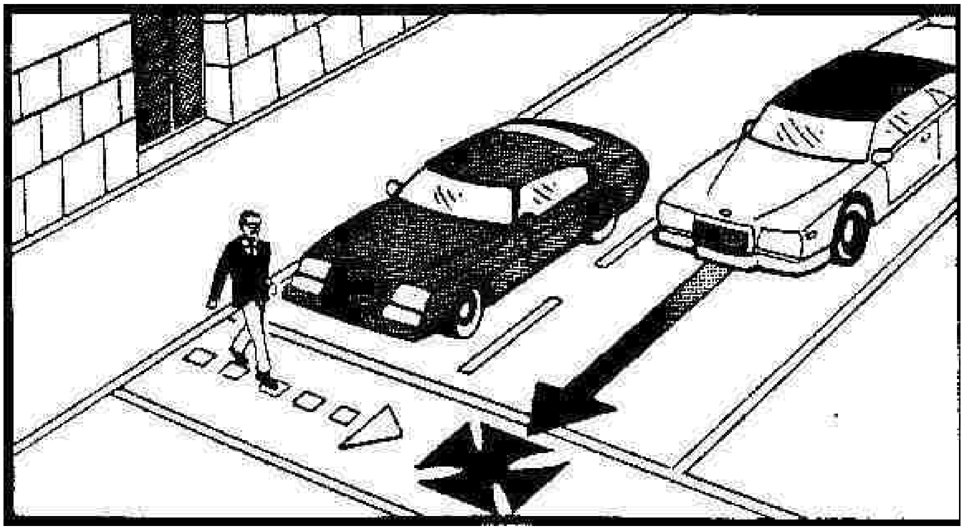


Figure 11. Illustration of partial-yield scenario

(Houten and Malenfant 2008; Zegeer et al. 2005)

On the other hand, partial-yield events involve situations on a multilane roadway, in which a vehicle in one lane stops to allow pedestrians to cross, while other vehicles in the adjacent lanes in the same direction do not stop for the same pedestrians (Houten and Malenfant 2008; Zegeer et al. 2005); the scenario is shown in Figure 11.

Of the two risky events, near-miss has been a topic of interest for a number of years, as it has been proven to be associated with crash occurrences (F. Guo et al. 2010). One of the earliest studies was performed in the early 1970s (Hayward 1972), and aimed to provide a better way of classifying near-miss events, since the classification was affected by subjectivity. The study suggested a threshold of one second time-measured-to-collision to be considered for classifying a maneuver as a near-miss event. Another notable early study on near-miss incidents was done in the mid-1980s (Houten, Malenfant, and Rolider 1985). This study was performed to assess the safety impact of posted feedback, a warning enforcement program, and pedestrians signaling before crossing the street, using a multiple baseline design (Houten, Malenfant, and Rolider 1985). The authors found that among the improved safety components, the near-miss incidents involving pedestrians declined by more than 50 %.

Since then, there have been a number of studies dedicated to near-miss events performed for various purposes using a variety of approaches (Matsui, Hitosugi, Doi, et al. 2013; Matsui, Hitosugi, Takahashi, et al. 2013; Sucha, Dostal, and Risser 2017; Voorhees 2017). In their study (Matsui, Hitosugi, Doi, et al. 2013) evaluated near-miss situations that involved car-to-pedestrians using pedestrian time-to-vehicle (pedestrian TTV,) which is considered as the time a pedestrian would require to reach the forward moving car line. The study used near-miss incidents recorded by cameras installed in different passenger cars in Japan. A total of 101 near-miss incidents were analyzed. They found that, on average for the near-miss incidents, the pedestrian TTV was about 1.05 seconds (Matsui, Hitosugi, Doi, et al. 2013), which was higher than that proposed by Hayward (Hayward 1972). Using the same data, another study (Matsui, Hitosugi, Takahashi, et al. 2013) was able to show the existing similarity between near-miss and real-world fatal crashes. Moreover, the study was able to determine the time-to-collision for car-pedestrian crashes, which was shorter under partial-yield scenarios than when driver-to-pedestrian view was unobstructed. This study proposed that automatic pedestrian detectors and braking systems to be installed in cars. One of the shortfalls of these studies is that they were not able to include other traffic conditions, crosswalk factors, or pedestrian characteristics in their analyses.

Different from the above mentioned studies that mostly used video cameras for data collection, the survey questionnaire and focus groups have been used (Voorhees 2017). This study aimed to understand challenging school crossings in New Jersey where near-miss incidents mostly occur. Police officers were tasked to respond to the survey, which aimed to collect a variety of information, including near-miss incidents. Out of 231 distributed surveys, 176 were returned, wherein 30% did not have challenging school crossing locations. A total of 186 challenging locations were identified, in which officers were aware of crashes and near-miss incidents at 21% to 81% of identified crossings. One of the criticisms of this study was subjectivity, as a "challenging location" could be defined differently by different officers. Thus, a better data collection technique is advised.

A combination of video cameras and survey questionnaires for data collection was applied (Sucha, Dostal, and Risser 2017) to evaluate the pedestrian-driver interaction at un-signalized marked crosswalks in urban areas in the Czech Republic. A total of 473 persons responded to the short interview, while 1584 observations were collected through video cameras. A logistic model was developed to associate the conflict situations, including near-miss events and other variables. The study found that in most cases, factors affecting pedestrian and driver actions and reactions were high vehicle pedestrian densities and vehicle speed. As this study was done on un-signalized crosswalks, some of their findings might not be applicable for signalized midblock crosswalks.

As partial-yield incidents are considered risky (Zegeer et al. 2005) to pedestrian safety, several studies (Fisher and Garay-Vega 2012; Garay-Vega 2008; M. F. Mitman, Ragland, and Zegeer 2008; Ragland and Mitman 2007; Zegeer et al. 2005) have been devoted to studying scenarios and resulting partial-yield crashes. A safety implications study of marked versus unmarked crosswalks at uncontrolled locations under various roadway conditions was performed in order to provide safer crossings for pedestrians (Zegeer et al. 2005). It used five-year crash data collected from 1,000 marked crosswalks, and 1,000 unmarked crosswalks with no traffic signals or stop signs on the approaches. It was revealed that a total of 17.6%, which was 33 out of 188 pedestrian crashes that occurred at marked crosswalks, were partial-yield crashes. On the other hand, none of the 41 pedestrian crashes at unmarked crosswalks was classified as a partial-yield crash. The authors provided two possible reasons for the high frequency of partial-yield crashes at marked crosswalks: one being a high likelihood of pedestrians stepping out in front of oncoming traffic at the marked crosswalk, while the second was that pedestrians are less likely to search properly for incoming vehicles before passing a stopped vehicle at marked crosswalks, compared to unmarked crosswalks. The study suggested detailed further research on the impact of an advanced yield line on pedestrian safety at marked crosswalks, which was later researched (Fisher and Garay-Vega 2012; Garay-Vega 2008). In their study (Fisher and Garay-Vega 2012; Garay-Vega 2008) found that an advance yield line leads to a change in driver behavior in terms of scanning for pedestrians and increased yielding distance, thus, improving pedestrian safety. Their study avoided subjecting staged pedestrians in partial-yield incidents due to the danger involved, and thus, used a driving simulator. Further, two studies (M. Mitman, Cooper, and DuBose 2010; M. F. Mitman, Ragland, and Zegeer 2008) concluded that partial-yield events were common on multilane roadways having four or more lanes with median refuge. However, no analysis was performed to link the number of lanes and partial-yield events.

It can be observed that previous studies have successfully associated near-misses and actual crashes but were unable to associate partial-yield and near-miss events. Therefore, it is still unknown to what extent partial-yield events can be risky to crossing pedestrians and drivers. Moreover, the factors associated with either partial-yield or near-miss incidents at signalized midblock crosswalks are yet to be explored. This study, therefore, focuses on exploring the relationship between partial-yield and near miss incidents; furthermore, it explores the factors associated with both partial-yield and near-miss events at signalized midblock crosswalks equipped with different types of signals. This study's findings and recommendations may be vital to traffic engineers to provide safe interactions between pedestrians and drivers at signalized crosswalks. If the associated factors for partial-yield and near-miss incidents are identified, the safety of crosswalks can be improved for the betterment of both pedestrians and drivers. Moreover, the study findings can be applied whenever a new crosswalk is to be installed.

2.3.3. Driver yielding dilemma

A yielding dilemma occurs when drivers do not understand what action to take when approaching activated flashing signals in the absence of pedestrians. This situation is common in PHBs, which in turn may result into unnecessary delays and crashes, as well. For instance, in the city of Lawrence, Kansas, a video-based observational study has shown that only 27% of the drivers took the correct actions when PHBs were flashing in the absence of pedestrians; the remaining 73% did not know what to do (Godavarthy and Russell 2016). In response to the situation, the city distributed handouts to drivers to educate them and solicit their understanding of PHBs, and a total of 35 completed surveys, out of 250 distributed surveys, were collected. The survey results depicted that most drivers responding to the survey only understood the blank signal phase (94%), while a relatively large percentage understood the steady red phase (91%). Only 15 out of 35 respondents understood the flashing yellow phase (Godavarthy and Russell 2016). Additionally, a driver's dilemma to yield to pedestrians might be influenced by the vehicles in the traffic stream, either on the adjacent lane or behind, as (Potts et al. 2015) observed that some drivers made last-minute decisions not to yield if the driver in the adjacent lane did not yield.

2.3.4. Yielding distance

As described earlier, drivers' yielding compliance at signalized midblock crosswalks is a safety assessment measure that has been extensively used for before-and-after countermeasure installation, as well as comparison of different crosswalk treatments (Brewer, Fitzpatrick, and Avelar 2015; Karkee, Nambisan, and Pulugurtha 2010; Kutela and Teng 2018; Nambisan et al. 2009). Drivers' yielding compliance is presented as the percentage of the incidents where drivers yielded right-of-way to pedestrians. For a before-and-after analysis, a higher yielding compliance after countermeasure installation indicates improved safety. Similary, any treatment that results into higher yielding compliance is indicative of a better safety performance, compared to other treatments. The literature, however, suggests that yielding compliance alone might not be a complete effectiveness measure, and the distance at which drivers yield right of way to pedestrians could convey more safety implication (Fitzpatrick et al. 2011). The yielding distance is defined as the distance between the crosswalk markings and the location that the incoming vehicle stops or reduces speed for pedestrians in the crosswalk (Nambisan et al. 2009). The longer the yielding distance, the more effective is the treatment.

Several studies have reported increased yielding distance after signalized crosswalk intallaments. A notable study (Nambisan et al. 2009) used 91and 116 observations before and after

RRFB installation, and found that the percentage of drivers yielding farther from the stripes increased significantly. The yielding distance was also associated with the number of RRFBs at the crosswalk in a study performed (Shurbutt et al. 2009). According to this study, the percentage of drivers who yielded at a distance of 100 ft or more doubled (7.2% to 15.1%) after RRFB installation. The sign and marking placement, in association with the yielding distance, was also evaluated (Houten, Malenfant, and McCusker 2001). Their study found that placing a yield markings and signs at 10m (33ft) before crosswalks produced similar benefits as placing them at 15m and 25m. In all cases, the presence of the yield markings and signs increased the yielding distance of drivers, which eventually reduced motorvehicle-pedestrian conflicts. Similar results were reported (Van Houten et al. 2002), who performed a study on streets with 50 km/h (30 mph) posted speed limits; the the presence of the signs and markings resulted in a low percentage of vehicle-pedestrian conflicts that involved evasive actions.

As mentioned earlier, most crosswalks have a yield line by which drivers are supposed to yield the right-of-way to pedestrians before crossing it. According to the manual on traffic control devices (MUTCD) (Manual on Uniform Traffic Control Devices (MUTCD) 2009), under normal traffic and land use conditions, the yield line should be positioned not less than 40ft and greater than 180ft from the crosswalk signal. Furthermore, under special scenarios, the yield line can be as close to the signal as possible. This study defines the spatial yielding compliance of drivers as the instance that a driver yields right-of-way to pedestrians before crossing the yield line. Although studies have attempted to evaluate the yielding compliance of drivers at varying distances from the crosswalk to yield line, two shotfalls are observed. First, studies have focused on a single factor (Houten, Malenfant, and McCusker 2001; Van Houten et al. 2002; Samuel et al. 2013); thus, neglecting the combined effects that other features might have. As a result of focusing on a single factor, descriptive analysis has been the dominant methodology for analysis.

2.3.5. Pedestrian and vehicle delays

Pedestrian and vehicle vehicle delays at signalized midblock crosswalks or intersections occur due to either of the following two reasons: either pedestrians are waiting for vehicles to yield, or drivers are waiting for pedestrians to cross (Nambisan et al. 2009). The magnitude of delays can be different depending on various factors including type of treatment. For instance, it is expected that a full signal would have higher average delays for both pedestrians and drivers, while a PHB would have high pedestrian delays, but low driver delays. On the other hand, RRFBs and CFBs are expected to have relatively low pedestrian and driver delays. A study (Nambisan et al. 2009) revealed that installation of an RRFB-based signalized midblock crosswalk resulted into an average of a 3.7 second reduction in pedestrian delay (7.5s to 3.8s,) while drivers experienced an average of a 1.5 second increased delay (0.5s to 2.2s). However, the difference was not statistically significant, at a 95% confidence level.

Further, a study (Hunter, Srinivasan, and Martell 2012) reported a relatively large average pedestrian delay of 10.1 seconds before installation of an RRFB and 5.2 seconds after installation. According to this study, the RRFB installation not only reduced the average delay, but also variations (standard devation) in pedestrian delays (15.6 to 6.2 seconds). This study recorded the longest delay of 89 seconds in the after period, while the before period’s longest delay was 40 seconds. Contrarily, (Foster, Monsere, and Carlos 2014) found that most pedestrians were able to use a crosswalk with very minimal delays if any; the maximum delays when the RRFBs were active and inactive were 15 and 20 seconds, respectively, which according to the authors, were in the acceptable levels of service (A, B and C) according the Highway Capacity Manual.

The PHBs were found to result into relatively higher pedestrian delays compared to other treatments. A before-and-after study (Eapen 2014) showed that, the average pedestrian delay for one week (9.55 sec) and one year (15.3 sec) after the installation of a PHB were higher than before installation (8.31 sec). However, the maximum delay before installation (131sec) was higher than that of one week (87 sec) and one year (95 sec) after installation.

2.4. Review of statistical models

In associating the relationships between various interractions at signalized midblock crosswalks, different statistical models have been developed. Binary-based, multinomial, and count-based models have been used for different purposes.

The yielding compliance in the driver-pedestrian interraction that has widely being modeled. Logistic regression has prodominantly been applied to model the yielding compliance of drivers, which has been considered to have two options: yield or not yield. The main advantage of utilizing logistic regression is that it permits the evaluation of individual crossing data rather than aggregated data (Fitzpatrick, Avelar, et al. 2016; Fitzpatrick, Brewer, and Avelar 2014). The application of logistic regression assumes that the logit transformation of yielding compliance (yield or not yield) is linearly related to predictor variables. With a linear relationship, the odds ratios are used to interpret the relatonships between yielding compliance and predictor variables. The odds are not directly related to yielding rate, but the probability is that motorists will yield given a predictor variable (Fitzpatrick, Avelar, et al. 2016; Fitzpatrick, Brewer, and Avelar 2014). The mixed effect logistic regression has been used to take care of the unobserved heterogeneity resulting from data clustering. The odds of drivers yielding when a crosswalk had RRFBs were found to be statistically significantly positively associated with the posted speed limit, in the city of Garland; however, they were negatively associated with the crossing distance, two way traffic, and in the city of Waco. On the other hand, when PHBs were considered, a higher posted speed limit and two way traffic were found to be negatively associated with the odds of drivers to yield, while longer crossing distances were positively associated with high yielding rates (Fitzpatrick, Avelar, et al. 2016; Fitzpatrick, Brewer, and Avelar 2014).

Recently, (Fitzpatrick, Brewer, et al. 2016) applied a count-based model (Negative Binomial) to associate yielding compliance to a number of predictor variables. The basis for their application of a Negative Binomial (NB) was the limitation in modeling yielding rates by using logistic regression. With logistic regression, the yielding rate depends on the number of yielding and non-yielding vehicles; for instance, if platoon of vehicles is observed at a crossing location on a roadway with two lanes in each direction, the number of yielding vehicles will be constrained to the number of lanes (two), while non-yielding vehicles are not constrained. The same scenario applies when there are few vehicles on the roadway; as a result, a platoon of vehicles tends to be associated with more non-yielding vehicles. Instead of binary choices (yield or not yield), this study used number of non-yielding vehicles as the response variables. According to this study, NB was the best option to model non-yielding vehicles at signalized crosswalks since the data structure resembes a negative binomial experiment. It should be noted that NB experiments involve counting the number of successes (non yielding vehivles for this case) until a predetrmined number of failures (yielding) occur. To take into account unobserved heterogeneity, the authors used a Negative Binomial Mixed-Effects Model (NBMEM). Their study found that ADT per lane, a 30mph speed limit, the presence of a transit stop within 200 ft, activation of only overhead RRFB, and far side direction were statistically significant, at a 90% confidence level, related to an increase in non-yielding vehicles. Concersely, the presence of a school within 0.5 mi, a 40mph speed limit, and a legend on the sign face were statistically significantly associated with low non-yielding vehicles. Other factors that were evaluated included the presence of a median refuge, the presence of supplementary signs at the crosswalks, and one-way or two-way traffic presence, which were not statistically significant, at a 90% confidence level.

The crossing compliance of pedestrians has been mainly modeled using two models: multinomial logit and logistic regression. (Brosseau et al. 2013; H. Guo et al. 2014; K. Kim, Made, and Yamashita 2008; Koh and Wong 2014; Rosenbloom 2009; Yanfeng et al. 2010; Zhou et al. 2013). Using multinomial logit, a study (Zhou et al. 2013) divided the temporal crossing behaviors into four categories: regular users, late starters, sneakers, and partial sneakers. The study found that arrival time, the presence of oncoming cars, and crosswalk length were the crucial factors for late starters, while gender and age were found to affect sneakers and partial sneakers, respectively. The multinomial logit model was also applied to study pedestrains' spatial crossing preferences (H. Guo et al. 2014), where overpass/underpass, crosswalk, and jaywalk were the three available choices. Their study found that safety, convenience, time saving, and additional distance due to detour were the main factors affecting proper use of a crosswalk. On the other hand, (Brosseau et al. 2013; K. Kim, Made, and Yamashita 2008; Koh and Wong 2014; Rosenbloom 2009; Yanfeng et al. 2010) applied logistic regression to study spatial and temporal crossing compliance.

Additionally, another study (K. Kim, Made, and Yamashita 2008) focused on the spatial crossing compliance of pedestrians at signalized and unsignalized intersections in Hawaii. In their study, pedestrians were observed to see if they used crosswalks or jawalked within 200 ft of the crosswalks. The study found that male pedestrians had higher odds of spatial crossing compliance violation, while children had a lower odds ratio. The same study found that hotel districts and residential areas were associated with higher spatial crossing compliance violations, but they were statistically insignificant. Similary, (Rosenbloom 2009) studied temporal crossing compliance, where pedestrians that arrived during red lights were observed. Their study found that male pedestrians and few people waiting at the curb were the most dominant factors for temporal crossing compliance.

Another study (Koh and Wong 2014) in Singapore evaluated pedestrian’s gape acceptance at signalized intersections. Their study found that longer available gap length, gap type, and cross stages were the focal factors for pedestrians’ gap acceptance. In addition to the aforementioned factors such as gender and age, a study revealed that both pedestrian wait time and intersection clearing time are associated with the violation of temporal crossing compliance (Brosseau et al. 2013). An attempt to jointly analyze spatiotemporal crossing compliance was performed by (Yanfeng et al. 2010). The study found that pedestrians’ age and number of campanions, attraction sites near the crosswalk, and crossing time have impact on spatiotemporal crossing compliance.

This study, therefore, proposes the use of multistate models (Luís Meira-Machado et al. 2009), which are in the family of hazard-based models, to associate yielding compliance and other covariates. The distinct advantage of the multistate model is not only its ability to model the partial-yield, which were not considered by models in the previous studies, but also the transition durations between state occurrences. To the best of the authors’ knowledge, no study has attempted to model the states’ transitions in yielding compliance by using multistate models.

Moreover, to this end, it can be observed that pedestrians’ crossing behaviors have extensively been studied; however, most of the studies considered either temporal or spatial crossing compliance separately. In reality, spatial and temporal compliance occur jointly (Yanfeng et al. 2010). Studies have applied two of the family of choice models, i.e multinomial logit and logistic regression, to model crossing compliance. However, considering joint spatiotemporal compliance, an additional choice based mode, such as ordinal models can also be applied. In the previous studies, the choice of the appropriate model type was not justified, as the performances of the applied models were not assessed. Therefore, this study aims to provide an assessment of the alternative models for the spatiotemporal crossing compliance of pedestrians at signalized midblock crossawalks. It evaluates three possible models and suggests the best performing model. Using the best performing model, the associated factors for spatiotemporal crossing compliance are evaluated.

CHAPTER 3: STUDY METHODOLOGY

This study seeks to perform various statistical analyses that will associate driver-pedestrian-infrastructure and the resulting risks at signalized midblock crosswalks. In so doing, the association of several crosswalk features, human factors, and traffic characteristics to driver-pedestrian and pedestrian-infrastructure interactions are performed, and various models are developed. To attain this study's objectives, different sets of crosswalks with varying quantities of features and traffic characteristics, as well as pedestrian demographics and activities are selected. The influence of these features, traffic characteristics, and pedestrian demographics and activities on drivers' and pedestrians' behaviors at the crosswalks is assessed through descriptive analysis and inferential statistics. For crosswalk features such as signals, different types of signals are selected, and a comparison of their influences is performed. For human factors, pedestrian actions before and during crossing the roadway at signalized crosswalks are observed. Moreover, drivers' actions in response to pedestrians' actions and crosswalk conditions at the times pedestrians want to cross or are crossing the roadways are recorded and analyzed.

3.1.Study site selections

Data is the integral part of any research; thus, obtaining proper data for a research study plays a vital role in the findings. When a portion of the population (sample) is studied, and results are to be generalized to the entire population, identifying relevant data collection sites is very important. Two methods, random sampling and purposive sampling, are commonly used to identify the samples to be included in a study (Cochran 1977). While the samples are chosen randomly under random sampling, the researcher needs to have a focus and prior knowledge of the observations of interest when applying purposive sampling. Oftentimes, such a knowledge is gained during research gap establishment, when performing the literature review.

As described earlier, this study aims at analyzing driver-pedestrian-infrastructure interactions to evaluate the roles of pedestrians, drivers, and facility features for safety analyses of the interrupted traffic flow at signalized midblock crosswalks; thus, the study sites (samples) are signalized midblock crosswalks. Signalized crosswalks have been installed in different areas across the United States and all over the globe. Due to the fact that standard criteria are used to decide upon the installation of signalized crosswalks, this study's findings can be transferable to other locations with similar characteristics if an appropriate sample size is selected. Therefore, purposive sampling was used to determine the number of study sites. Several factors were considered in the selection of the study sites, so that the study sample is as inclusive of all communities as possible for findings transferability. The following criteria were used to determine number of samples.

3.1.1. Crosswalk characteristics

A number of crosswalk characteristics were considered, including the geometry of the crosswalks (Danish offset and straight crossing); signal types, which included the yellow flashing signals (RRFBs, CFBs, and CRFBs) and those displaying solid red (TCSs, and PHBs); crossing stages (one and two crossing stages); signal and pushbutton locations (sideways, median, and overhead); supplementary signage ("YIELD HERE TO PEDESTRIANS," "USE CROSSWALK," yellow and green "PED XING," "LOOK BEFORE CROSSING," and "STATE LAW YIELD TO PEDESTRIANS"); and pushbutton type (traditional, audible, and illuminated). In addition, the variations of the distances between the marked stripes and yield lines, as well as the yield lines and advanced pedestrian crossing signs (APCS) were involved in site selection.

3.1.2. Roadway characteristics

The geometry of the roadway on which the crosswalk is located also a focus. The important features were the presence of a median refuge, type of median (if present), and number of lanes, as well as the presence of exclusive bus lanes and turning lanes. Apart from geometry, vehicular traffic is the key observation in this study. The aspects of vehicular traffic considered were traffic volume and traffic speed. Traffic volume was presented as AADT, while traffic speed was estimated based on speed limit.

3.1.3. Land use characteristics

Sites were also selected based on the land use where they are located. Crosswalks located in several land use areas ranging from pure residential and pure commercial, to a mixture of residential and commercial were selected. Other special land use, such as University and school zones, were also considered.

3.1.4. Traffic crash history

Traffic crash history can have an influence on both pedestrians' and drivers' behaviors. Since most pedestrians are local to a particular area where a crosswalk is located, it is assumed that they are aware of the traffic crash history. This awareness could alter pedestrian behavior, which could eventually affect drivers' reactions. Thus, crosswalk locations that had histories of severe pedestrian crash occurrences (fatalities), injuries to pedestrians and vehicle occupants, and Property Damage Only (PDO) of vehicles, either before or after crosswalk installations, were considered.

3.1.5. Demographic characteristics

The demographic characteristics of the populations where the crosswalks are located was also included in site selection. The census tract and census block level population size, which in turn may translate into pedestrian volume, were the focus. Additionally, population distribution by race/ethnicity within the census tract where the crosswalks are located was also considered. The level of income in the zip codes where the crosswalks are located could influence the pedestrian volume, and thus, was another of the criteria for site selection.

Based on the criteria mentioned above, a sample of 20 signalized midblock crosswalks located in Southern Nevada were selected for analysis. Figure 12 shows the spatial distributions of the sites, while Table 1 shows the key characteristics of the sites.

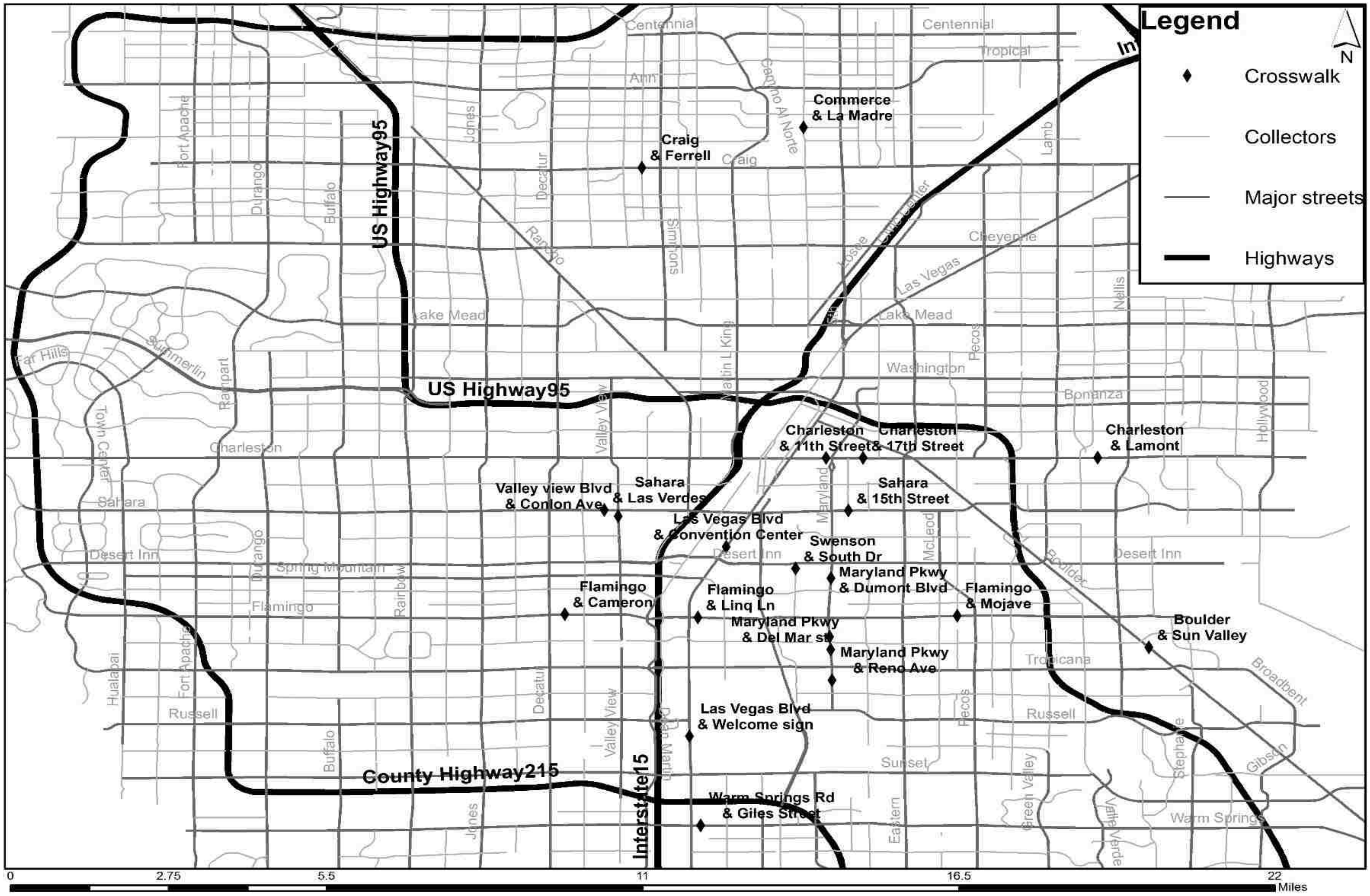


Figure 12. Spatial distribution of study sites

Table 1. Characteristics of data collection sites

sn	Main street	Minor street	Signal Type	Geometric Configuration	Crossing type	Land use	Speed limit	# Lanes	Block pop	2016 AADT	Median type
1	Boulder	Sun Valley	RRFB	Danish offset	Two Stage	Residential-commercial	45	Nine	3,870	34,000	Raised & wide
2	Charleston	11th Street	RRFB	Danish offset	Two Stage	Residential-commercial	35	Six	1,851	33,000	Raised & narrow
3	Charleston	17th Street	RRFB	Danish offset	One/Two Stage	Residential	35	Seven	2,676	33,000	Raised & narrow
4	Charleston	Lamont	RRFB	Danish offset	One/Two Stage	Residential	45	Seven	3,188	53,000	Raised & narrow
5	Commerce	La Madre	RRFB	Straight	One stage	Residential	30	Five	2,938	14,000	TWLTL
6	Craig	Ferrell	RRFB	Straight	One stage	Residential	45	Ten	1,643	31,000	Raised & narrow
7	Flamingo	Cameron	RRFB	Straight	One stage	Residential-commercial	45	Eight	4,029	51,000	Raised & narrow
8	Flamingo	Linq Ln	TCS	Straight	One stage	Commercial	35	Six	124	49,000	Raised & narrow
9	Flamingo	Mojave	RRFB	Danish offset	One stage	Residential-commercial	45	Eight	2,563	42,000	Raised & narrow
10	Las Vegas Blvd	Convention Center	TCS	Straight	One/Two Stage	Commercial	45	Six	1,514	38,000	Raised & wide
11	Las Vegas Blvd	Welcome sign	TCS	Danish offset	Two Stage	Commercial	45	Six		47,000	Raised & wide
12	Maryland Pkwy	Del Mar st	TCS	Straight	One/Two Stage	University, residential-commercial	30	Eight	3,936	29,000	Raised & narrow
13	Maryland Pkwy	Dumont Blvd	CFB	Danish offset	One stage	Residential-commercial	30	Six	3,584	36,000	TWLTL
14	Maryland Pkwy	University Ave	CFB	Danish offset	One stage	University-residential	30	Seven	3,433	29,000	Raised & narrow
15	Maryland Pkwy	Reno Ave	CFB	Danish offset	One stage	Residential-commercial	30	Five	5,940	21,000	Raised & narrow
16	Sahara	15th Street	PHB	Straight	One/Two Stage	Residential	45	Eight	4,505	44,000	Raised & wide
17	Sahara	Las Verdes	RRFB	Straight	One/Two Stage	Residential-commercial	45	Nine	3,546	65,000	Raised & narrow
18	Swenson	South Dr	CFB	Straight	One stage	Commercial	30	Five	6,680	13,000	TWLTL
19	Valley view Blvd	Conlon Ave	CRFB	Straight	One stage	Residential	30	Six	2,896	28,000	Raised & narrow
20	Warm Springs Rd	Giles Street	CFB	Straight	One stage	Commercial	45	Seven	432	26,000	TWLTL

3.2. Sample size estimation

Although purposive sampling was used for the selection of the number of study sites, the estimation of the number of observations from each site is based on random sampling. This is due to the fact that, although the number of study sites is fixed, the selected observations within each site are random, and for this case it comes from a finite population. The sample size estimation procedure developed by (Cochran 1977) was applied. According to (Cochran 1977), to determine sample size from a finite population, the critical value (Z) of the desired confidence interval, the estimated proportion of an attribute that is present in the population (p), and the desired level of precision (C) should be specified. It should be noted that $(q = 1 - p)$. The sample is first assumed to be from an infinite population; then a population correction factor is applied for a finite population. Therefore, for infinite population, the minimum sample size is estimated as (equation 1):

$$n_r = \frac{Z^2 * p * q}{C^2} \tag{1}$$

Since the variability is not known, the maximum variability, which equals to 50%, is assumed $(p = 0.5)$. Furthermore, assuming a precision $\pm 5\%$, and the confidence level considered to be 95% (which implies a Z-score of 1.96), the minimum sample size for each study site is estimated to be:

$$Sample\ size = \frac{1.96^2 * 0.5 * 0.5}{0.05^2} = 384\ observations \tag{2}$$

Further assumptions are made to obtain the proportion of pedestrians from the block and census tract population. It is assumed that people who live within 0.25 miles of the crosswalk are more likely to use the crosswalk. Therefore, using Geographical Information System (GIS), a buffer of a 0.25-mile radius was created on a block population shape file, and the sum of the people living within 0.25 miles of each crosswalk was determined (Table 1). However, only a portion of the population is pedestrian. Estimating pedestrian flow, the findings from the Southern Nevada Household Travel Survey performed in 2014 (RTCSN 2015) was used. The (RTCSN 2015) found that on average, 8.3% of the households in Southern Nevada have no vehicles; the statistics, however, are higher for Clark County – Paradise (17.5%) and East Las Vegas (16.8%) , and lower for County Unincorporated (1.3%), Clark County – Southwest (3.6%), and the City of Henderson (3.4%).

As expected, the jurisdictions with high percentages of households without vehicles have high percentages of households without licensed drivers and have high percentages of non-motorized trips. In fact, the regional average of non-motorized trips is 12%, while East Las Vegas (18%) and Clark County – Paradise (19%) have the highest percentages, compared to the regional average. Therefore, it can be assumed that, pedestrians and bicyclists account for 12% of all of the population across the region, but the percentage is relatively higher for the locations with high non-motorized trips. Therefore, for such locations, 20% of the population is assumed to be pedestrians and bicyclists. Applying a population correction factor for equation 2, the sample size for each site can be computed as:

$$n = \frac{n_r}{1 + \frac{n_r - 1}{N}} \tag{3}$$

where n is the estimated minimum sample size, n_r is the sample size for an infinite population, and N is the population size. The minimum sample size varies from 100 observations to 300 observations.

3.3. Data collection procedure

Data collection was performed by using a video camera, which was positioned in such a way that pedestrians and drivers did not easily detect it, in order for them to maintain their natural behaviors. The camera was also positioned to capture as much information as possible occurring within the "effective crosswalk distance." The procedure used to define the "effective crosswalk distance" is similar to the one presented by (Sisiopiku and Akin 2003); however, contrary to their study, for this study there are clear demarcated distances. The effective crosswalk distance (Figure 13) is considered as the entire distance upstream and downstream of the marked stripes, bordered by the advanced pedestrian crossing signs (APCSs).

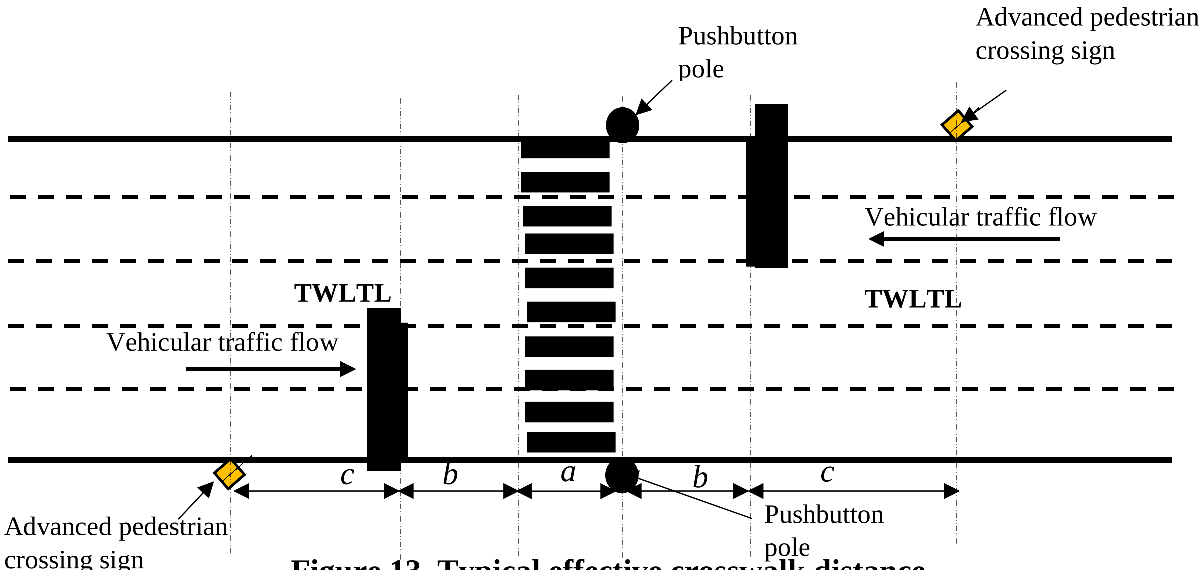


Figure 13. Typical effective crosswalk distance

Here, "a" is the marked/stripped zone of the crosswalk, "b" is the distance between the marked strips and yield line, "c" is the distance between the yield line and advanced pedestrian crossing sign, and "$a + b + c$" is the effective crosswalk distance (zones dimensions vary per crosswalk).

Within this distance, crossing activities are assumed to be influenced by the existence of the crosswalk, and drivers are assumed to be aware of the possibility of the presence of pedestrians. The effective crosswalk distance can be segmented into three distinct zones: “zone a,” which is within the marked stripes; “zone b,” which is between the marked stripes and the yield line; and “zone c,” which is within yield line and advanced pedestrian crossing sign (Figure 13). These zones are important in identifying and analyzing pedestrian crossing compliance, since according to the state law, pedestrians are supposed to use the striped zone to cross the roadway. Since crosswalks are not homogeneous, the effective crosswalk distance is not constant; it varies per crosswalk. Therefore, the dimensions of zones “a,” “b,” and “c” are not fixed. The typical APCSs, and the distances between them to the marked areas of the crosswalks are shown in Figure 14.

Figure 14. Advanced pedestrian crossing signs with and without flashers

It can be observed that some crosswalks have flashing signals at the APCSs. On these crosswalks, drivers become aware of the pedestrians' occupancy or intention to use crosswalks far in advance, due to the presence of flashing lights.

Video recording was performed on an hourly basis for easily analysis of vehicle and pedestrian flow (i.e. pedestrians per hour and vehicles per hour). During data collection, the specific times of the day were considered, which are, morning (time to go to work/school), afternoon (lunch time), and evening (time to go back home from school/work). Additionally, data were collected for both day and night times for weekdays and weekends in order to capture the most variabilities. At least three hours of data collection were performed at each crosswalk. For some crosswalks with low pedestrian intensities, more hours of data collection were assigned.

3.4.Data processing

Data processing involved the extraction of observations from videos to an excel spreadsheet. This was performed through watching videos and extracting the observed behaviors of pedestrians and drivers. The following information from each crossing incidence were of interest: First, pedestrians were observed to determine whether they press pushbutton before crossing. Then, pedestrian crossing zones and the yielding behavior of the drivers were extracted and recorded. To preserve the natural crossing behavior when extracting the crossing patterns of pedestrians, both the starting and ending zones where pedestrians crossed through were considered. This is to say, if a pedestrian started crossing between the marked stripes and yield line (zone b), and finished crossing between the yield line and the advanced pedestrian crosswalk signs (zone c), the crossing incident was recorded as occurring in two zones (zone b and c). Moreover, as pedestrians pressed the button or stood at the curb/sidewalk, which indicated an intention of using the crosswalk, time for the initial state (non-yield) began when the first non-yielding vehicle passed the crosswalk. If a vehicle in any lane stopped to allow pedestrians to use the crosswalk, then the time at which the vehicle stopped was recorded. If the stoppage involved all vehicles, it was recorded as a full-yield state; however, if vehicles in one or more lanes did not stop, the resulting state was recorded as "partial-yield," which implies that the time continued to be recorded until the full-yield occurred. Other potential variables of interest which are described in the next section were also extracted.

3.5.Potential Variables

As mentioned earlier, this study is interested in analyzing the interactions between pedestrians, drivers, and crosswalk features under interrupted traffic flow at signalized midblock crosswalks. Key observations that express the roles of all participants in the interrupted traffic flow at the crosswalks are:

i. Drivers yielding compliance, which was determined by considering not only whether drivers yielded right of way, but also the time taken for them to yield, the zone in which they yielded, and whether all the drivers yielded, or if a partial-yield state was observed. Therefore, the time from when the first non-yielding vehicle passes waiting pedestrians to the time all the vehicles completely yield is crucial.

ii. Presence of near-crash events, which were defined as situations where a vehicle was about to get involved in a crash with either a pedestrian or another vehicle as a result of the interrupted traffic flow by a pedestrian.

iii. Pedestrian compliance to use pushbuttons, and their spatial and temporal crossing compliance in response to the signs and directives provided at the crosswalk. These signs and directives include the pushbutton sign that was mostly provided at the pushbutton; look before crossing sign, which was often placed either at the side poles or at the median; and use crosswalk signs which were mostly positioned at the median, especially for the crosswalks whose pedestrians have a history of jaywalking. Therefore, pedestrians were observed to determine whether they pressed the pushbutton before entering the crosswalk, looked for incoming traffic before crossing, and crossed within the dedicated crossing zones. The latter facilitates in determining the pedestrian spatial crossing compliance.

The association of the aforementioned variables to the explanatory variables can be determined by the statistical models (inferential statistics).

3.5.1. Potential dependent variables

The potential dependent variables can be categorized into two categories: traffic related and pedestrian related. Traffic related variables include the temporal and situational variables, whereby the temporal variables involve the time taken for the transitions from non-yield states to full-yield states, non-yield states to partial-yield states, and partial-yield states to full-yield states. In a similar fashion, during the same transitions, the numbers of vehicles were counted. Situational variables for this case include full-yield state, partial-yield state, and near crash event, which describe whether the full-yield, partial-yield, and near crash events were observed, and spatial yield compliance, which represents the situation in which the vehicle(s) stopped before the yield line. Pedestrian-related variables mainly focused on understanding whether pedestrians used the pushbutton and complied with spatial and temporal crossing requirements. Thus, the pressed variable was assigned to a scenario in which a pedestrian pressed the button, temporal crossing compliance meant that a pedestrian waited for the walk signals at TCSs and PHBs, and spatial crossing compliance was assigned when a pedestrian crossed within marked stripes. Spatiotemporal crossing compliance considered that both spatial crossing compliance and temporal crossing compliance occurred jointly.

3.5.2. Potential explanatory variables

Several variables may affect pedestrians' behaviors towards effectively utilizing signalized crossing locations, as well as the yielding behavior of drivers. These variables are grouped into four groups: crosswalk related, pedestrian related, traffic related, and temporal related variables.

3.5.2.1.Crosswalk related variables

Crosswalk related variables in this study include a variety of crosswalk features and characteristics. The first is the signal type at the crosswalk, whereby the influence of CFB, CRFB, PHB, TCS, and RRFB signal types on driver and pedestrian behaviors are evaluated. Secondly, the crosswalk geometry, which is defined in terms of cross stages, median type, yield line to marked stripes distance, yield line to APCS distance, and the presence of turn lanes were also linked to pedestrians' and drivers' behaviors. The cross stage is described as the number of stages that pedestrian needs to go through when crossing the roadway. One cross stage means the crosswalk is designed in such a way that a pedestrian will start and finish crossing without stopping at the middle. In other words, no median refuge is provided for pedestrians. With two stages, on the other hand, pedestrians need to stop at the median refuge before crossing the second side of the road. Crash history is another factor that could especially affect pedestrian crossing behavior. The assumption is that since the same pedestrians regularly utilize a crosswalk, for them, knowing that the crosswalk is prone to crashes would improve their crossing compliance. The number of lanes is expected to have a significant effect on the use of a pushbutton. Land use where the crosswalk is located is also of the interest. Three types of land use, mixed, commercial, and residential, were identified. The presence of signs such as yellow "PED XING," "YIELD HERE TO PEDESTRIANS," "USE CROSSWALK," "PUSH BUTTON TO TURN ON LIGHTS," and "STATE'S LAW YIELD TO PEDESTRIAN" are also of interest. The crosswalk signalization status, as well as the nearby traffic signal status, are also included in the crosswalk related variables to be used in the analysis.

3.5.2.2.Pedestrians related variables

Pedestrian based variables include age, gender, arrival sequence, pedestrian crossing zone, activities before and during crossing, number of pedestrians in one crossing incident, and whether a pedestrian was either coming or going to a bus. Since pedestrians were not asked about their age, this variable was approximated based on visual judgement. Approximate pedestrians' ages were categorized into five groups: children and teens, young adults, adults, and elderly. The mixed age group was assigned when pedestrians with different age groups crossed together. In a similar fashion, the race variable was determined. This variable was categorized into these categories: White or Hispanic only, Black only, and mixed races, where different races used a crosswalk at the same incidence. Different activities before and during crossing ware also observed. The identified activities include holding/carrying things, pushing things (bag, stroller, or cart), riding a bike, and using a phone. The pedestrians' crossing zones were identified according to the zones in which they started and finalized their crossings

3.5.2.3.Traffic related variables

The traffic related variables in this study include the annual average daily traffic (AADT), number of vehicles within effective crosswalk distance (ECD), incoming vehicle speed, and vehicle position when pedestrian arrives at a crosswalk. The number of vehicles was determined by counting the number of vehicles in both directions of travel within the effective crosswalk distance. The incoming vehicle speed was estimated by dividing the fixed distance and time taken to cross that distance. The distance was measured on-site, while the time was determined by watching the video and using a stopwatch to estimate the time taken to cross the measured distance. The vehicle position variable was determined by considering the position of the front vehicle when a pedestrian arrived at the crosswalk. This was determined first by recording the zone in which the vehicle was located; then these zones were converted to distances.

3.5.2.4.Temporal related variables

The only temporal variable in this study is the time of the day. This is the variable representing the time of the day when the data were recorded. The time of the day was divided into morning (7:00am-11:00am), early afternoon (11:00am-1:00pm), late afternoon (1:00pm-4:00pm), evening (4:00pm-6:00pm), and night (6:00pm-9:00pm).

This study seeks to perform various statistical analyses that can associate several interactions occurring at signalized midblock crosswalks with their explanatory variables. In so doing, analyses of several crosswalk features, human factors, and traffic characteristics in connection to driver-pedestrian and pedestrian-infrastructure interactions are performed, and various models are developed. To attain this study’s objectives, different sets of crosswalks with varying quantities of features and traffic characteristics, as well as pedestrian demographics and activities are selected. The influence of these features, traffic characteristics, and pedestrian demographics and activities on drivers’ and pedestrians’ behaviors at the crosswalks is assessed through descriptive analysis and inferential statistics. For crosswalk features such as signals, different types of signals are selected, and a comparison of their influences is performed. For human factors, pedestrian actions before and during crossing the roadway at signalized crosswalks are observed. Moreover, drivers’ actions in response to pedestrians’ actions, and crosswalk conditions at the time pedestrians want to cross or are crossing the roadway are recorded and analyzed.

3.6.Statistical model development

This study aims at developing two types of models, hazard-based and choice-based models, to associate the dependent and explanatory variables of interest. Hazard-based models, specifically multistate models, evaluate not only the probability of event occurrence, but also the relative time taken for that event to occur. The models are applied in evaluating the factors that affect the transitions from one state to another during the yielding compliance of drivers. The second model types to be developed are the choice-based models, which are basically logistic regressions, multinomial logistic regressions, and ordered logistic regressions. These models are developed to assess the pushbutton use, pedestrian crossing compliance, near crash event occurrence, pedestrian spatiotemporal crossing compliance, and driver spatial yielding compliance.

3.6.1. Hazard-based multistate models

Contrary to previous studies that modeled yielding compliance by using either choice-based models (Fitzpatrick, Avelar, et al. 2016; Fitzpatrick, Brewer, and Avelar 2014) or frequency-based models (Fitzpatrick, Brewer, et al. 2016), this study presents the yielding compliance by using a hazard-based model. The reason (Fitzpatrick, Brewer, et al. 2016) opted for a frequency based model instead of a binary choice model is the fact that the number of vehicles yielded is constrained by the number of lanes, while the number of non-yielding vehicles is never constrained. For instance, for the case in which there is platoon of vehicles on a three-lane roadway, regardless of the number of non-yielding vehicles, the number of yielding vehicles will remain constrained to three. Therefore, using the number of non-yielding vehicles as a measure of yielding compliance was the best option (Fitzpatrick, Brewer, et al. 2016). However, the use of the number of non-yielding vehicles may be deceiving; this is, when counting the non-yielding vehicles, no consideration of the vehicles arrangements is given. For instance, for a three-lane roadway, three vehicles may pass the crosswalk location either in parallel or in series. The time spent by the three vehicles in series to pass the crosswalk location is quite different from the time spent to pass the same location if vehicles are in parallel; however, the frequency-based models would still use the same number of vehicles. Hence, time until vehicle yield would be a consistent measure of yielding compliance. To model the actual time until vehicle yield, hazard-based models are applied.

3.6.1.1.Concept

Hazard-based models are commonly used to associate the explanatory variables to dependent variables, when modeling the time to event occurrence is the target. The models have been extensively applied in the medical field, where time to occurrence of particular event such as death or disease is to be investigated (Giard, Lichtenstein, and Yashin 2002; Hougaard 1999; Luís Meira-Machado et al. 2009). In recent years, hazard-based models have been applied in traffic engineering to understand incident durations and traffic patterns (Chimba et al. 2014; J. Kim, Mahmassani, and Dong 2010; Laflamme and Ossenbruggen 2017; Nam and Mannering 2000). However, to the best knowledge of the author, no attempt to apply hazard-based models, multistate models in particular, to explain yielding compliance of drivers.

Before occurrence of an event of interest, transitions between different states are possible. In most cases, three types of states are observed, which are: initial states, where a subject enters the study area; absorbing states, which is the endpoint of the study; and transient states, which comprise all intermediate states. For instance, in the medical field, the states can be conditions such as healthy, diseased, and dead, while transitions can be disease outbreak or death occurrence (Hougaard 1999). The full statistical model specifies the state as well as the transitions from state to state in the form of the hazard function, also known as the intensity function $\lambda(t)$. For events that involve transitions between states, traditional hazard-based models fail to consider the states involved during any event of interest. When the time-to-event is discretized into a distinct state, the multi-state models are the best preference (Giard, Lichtenstein, and Yashin 2002; Hougaard 1999; Luís Meira-Machado et al. 2009). Consider a two-state model (Figure 15) in which only one transition $\lambda(t)$ is observed. A person can be healthy and transition to death; a typical example is a person died due to a heart attack.

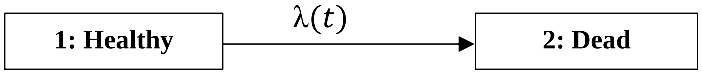


Figure 15. The two-state model for survival data

(Luís Meira-Machado et al. 2009)

Moreover, multiple state models, whereby three or more states are observed, are also possible. The two typical examples are: the illness–death or disability model (Figure 16) and the permanent illness–death model (Figure 17) (Luís Meira-Machado et al. 2009).

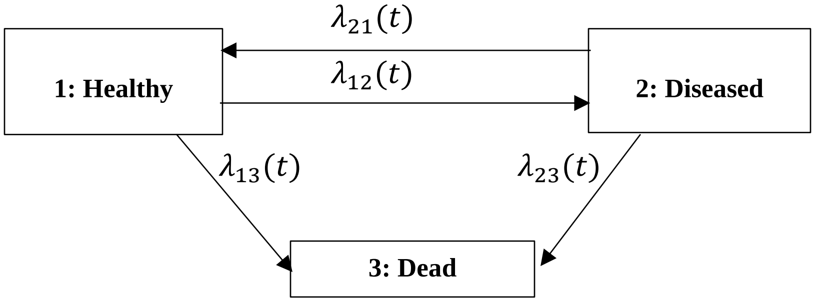


Figure 16. Disability model

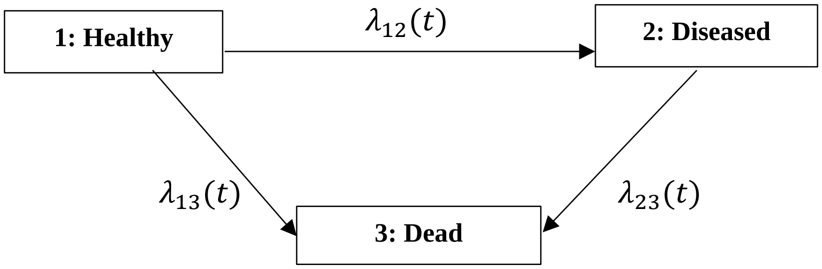


Figure 17. Permanent illness-death model
(Luís Meira-Machado et al. 2009)

In the disability model, a person can be disease free, then contact certain disease, and either get back into the disease-free state or move to the death state; or a person can be healthy and move directly to the death state. On the other hand, for the permanent illness-death model (Figure 17), the transition from diseased to disease-free is not possible.

In the context of yielding compliance, when a pedestrian encounter flowing vehicular traffic at a signalized crosswalk, three states and three transitions are possible: a non-yield state, partial-yield state, and full-yield state. This is, after certain time (T), whereby T varies from 0 to t, vehicles may either continue to flow, a situation described in this study as a non-yield” state, or transit directly into a full-yield state. The other option is transiting into a full-yield state through a partial-yield state. The non-yield state is when all vehicles continue to flow without stopping or yielding; the partial-yield state is when vehicles in one or more lanes yield, while other lanes continue flowing; whereas the full-yield state occurs when vehicles in all lanes yield or come to a complete stop. Diagrammatically, these states and their corresponding transitions can be presented as shown in Figure 18. The states and transitions in driver yielding compliance are similar to the permanent illness-death model shown in Figure 17.

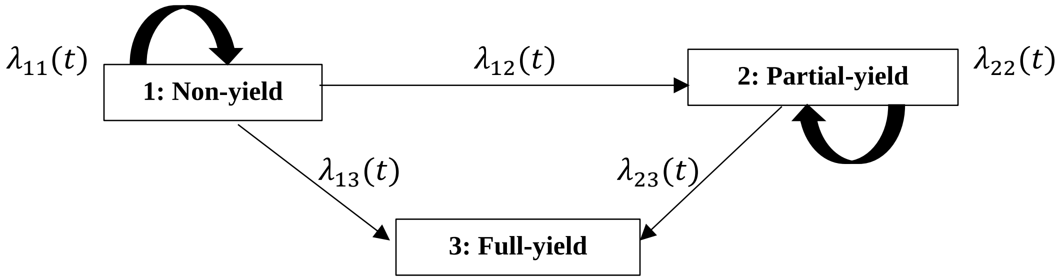


Figure 18. Drivers yielding states and possible transitions

3.6.1.2.Multistate model estimation

The parameters of interest in multistate models are: (i) the relationship between covariates and time to event; (ii) the transition intensities (hazard rates); (iii) the transition probabilities; (iv) the state occupation probabilities; and (v) the distribution of time spent in each state (Araújo, Meira-Machado, and Roca-Pardiñas 2014; Kneib and Hennerfeind 2008).

As described above, driver yielding can occur either just after a pedestrian has shown the desire to cross the roadway ($time\ T = 0$) or after certain period ($time\ T = t$). The driver yielding compliance for this case consists of a random variable, T, which represents the time elapsed until part or all of the drivers in the traffic stream yield ($T \leq 0 \leq t$). Distribution of time (T) is characterized by the survival function $S(t) = P_r(T > t)$ or the transition intensities (hazard rate) $\lambda(t)$ given as:

$$\lambda(t) = \frac{\partial \ln(S(t))}{\partial t} = \lim_{\Delta t \to 0} \frac{P_r(t \leq T < t + \Delta t | T \geq t)}{\Delta t} \tag{4}$$

As shown in the Figures 17 and 18 above, in a multi-state model, at any time there is a state that is occupied, and there exists a probability of transition between states. According to (Machado 2011), the transition probabilities between states, h and j for $s \leq t$ is given as:

$$p_{hj}(s,t) = p(X(t) = j | X(s) = h, H_{s-}) \qquad (5)$$

In equation 4, H_{s-} denotes the history of the process, which consists of observation over either the interval of time bounded by $(0, s)$, or through transition intensities representing instantaneous hazards of progression from state h to state j.

Let the underlying stochastic process be denoted by $\{X(t), t \geq 0, X(0) = 1\}$, whereby $X(t)$ represents the occupied state at time t (all individuals are in state 1 when time equals zero). If T_{hj} represents the possible transition from state h to state j, for this case ($1 \leq h \leq j \leq 3$), then the stochastic behavior of the process can be represented by a random vector of transition time (T_{12}, T_{13}, T_{23}). Therefore, the survival time is given as:

$$T = I(T_{12} \leq T_{13})(T_{12} + T_{23}) + I(T_{12} > T_{13}) + T_{13} \qquad (6)$$

Not all the time in the event of interest is observed; the presence of such a situation introduces the so-called right-censored observations at time t. Thus, the right-censoring variable C, which is assumed to be independent of (T_{12}, T_{13}, T_{23}), is introduced. With the right-censored parameter introduced, the possible sojourn time U in state 1 can be $U = \min(T_{12}; T_{13}; C)$; the sojourn time V in state 2 can be $V = \min(T_{23}; C - T_{12})$; whereas the observed total time Y is given as $Y = U + \delta V$; this is, $Y = \min(T, C)(\delta = I\,(T_{12} \leq \min(T_{13}; C)))$; and indicator statuses $\Delta_1 = I(\min(T_{12}; T_{13} \leq C)$ and $\Delta_2 = I(T \leq C)$. Following additive probability rules, the following relationship can be established:

$$p_{13}(s,t) = 1 - p_{11}(s,t) - p_{12}(s,t) \qquad (7)$$

$$p_{23}(s,t) = 1 - p_{22}(s,t) \qquad (8)$$

Therefore, upon determining three transition probabilities, probabilities $p_{11}(s,t)$, $p_{12}(s,t)$, and $p_{22}(s,t)$, the remaining two transition probabilities, $p_{13}(s,t)$ and $p_{23}(s,t)$ can be obtained by using additive probability rules. The first three probabilities can be estimated as Markovian or non-Markovian. Under the Markovian assumption, also known as memory less, the future state is independent of the past states, but dependent of the current state only (Luís Meira-Machado, de Uña-Álvarez, and Cadarso-Suárez 2006). According to (Datta and Satten 2001; Luís Meira-Machado, de Uña-Álvarez, and Cadarso-Suárez 2006), the Aalen-Johansen estimator (Aalen and Johansen 1978) is suitable for a non-Markovian process when the target is in occupancy probabilities not transition probabilities. (Luís Meira-Machado, de Uña-Álvarez, and Cadarso-Suárez 2006) proposed the use of the Kaplan-Meier estimator for transition probability estimations under a non-Markovian situation. According to (Luís Meira-Machado and Roca-Pardiñas 2011) the proposed Kaplan-Meier estimator for transition probabilities from state h to state j are given as:

$$\hat{p}_{11}(s,t) = \frac{1-\widehat{H}(t)}{1-\widehat{H}(s)} \tag{9}$$

$$\hat{p}_{12}(s,t) = \frac{\sum_{i=1}^{n} W_{i\emptyset s,t}(U_{[i]}, Y_{(i)}}{1-\widehat{H}(s)} \tag{10}$$

$$\hat{p}_{22}(s,t) = \frac{\sum_{i=1}^{n} W_{i\emptyset s,t}(U_{[i]}, Y_{(i)}}{\sum_{i=1}^{n} W_{i\emptyset s,s}(U_{[i]}, Y_{(i)}} \tag{11}$$

Whereby W_i denotes Kaplan-Meier weights associated to $Y_{(i)}$, $\widehat{H}$ is the Kaplan-Meier estimator based on the pairs(U_i, Δ_{1i}), $\emptyset_{s,t}(u,v) = I(s < u \le t, v > t)$, and $\emptyset_{s,t}(u,v) = I(u \le s, v > t)$. The ultimate focus is not only determining the transition probabilities between states, but rather associating the outcome and explanatory variables (Z). The models assume the outcome variable is the linear function of the predictor variables, with unknown regression coefficients βs that are to be estimated (Andersen, Maja, and Perme 2008). This is:

$$LP(t) = \sum_{h=1}^{j} \beta_{hj}(t) Z_{hji}\,(t) \tag{12}$$

A multiplicative link function, i.e. $log(\lambda(t; Z))$, that relates hazard rates $\lambda(t; Z)$ and linear covariates $LP(t)$ has always been preferred (Martinussen and Scheike 2006). Upon fitting transition intensities that include explanatory variables in the non-parametric model above, and not specifying the baseline hazard λ_{hj0}, the resulting semi-parametric Cox regression (Andersen, Maja, and Perme 2008), also known as Cox Markovian models (CMM) since it assumes Markovian process holds, is given as:

$$\lambda_{hj}(t|Z) = \lambda_{hj0}(t) exp \sum_{h=1}^{j} \beta_{hj}(t) Z_{hji}\,(t) \tag{13}$$

However, as per Markovian assumptions, the future state depends only on the current state, not on the history. This assumption may or may not hold in the yielding compliance of drivers. Therefore, some modifications in Equation 13 are deemed. The best alternative is the use of Cox semi-Markovian models (CSMM), also known as “clock reset,” by which, the future state depends not only on the current state, but also current duration. This is, the hazard at time t depends on both: the state at time t, and duration $t - T$ at that state, where $t \leq T$. The new equation can be written as (Andersen, Maja, and Perme 2008):

$$\lambda_{hj}(t|Z) = \lambda_{hj0}(t) ex\,p(\beta_0 f(t - T) + \sum_{h=1}^{j} \beta_{hj}(t) Z_{hji}\,(t)) \tag{14}$$

This model is semi-Markov, if and only if β_0 = 0; such a condition enables testing the Markov hypothesis. The coefficients are estimated by using the maximum likelihood method.

The likelihood function for model parameters is expressed via Jacod’s formula as:

$$L(\Theta) = \prod_{i} \prod_{h} \prod_{j} \left(\prod_{t} \lambda_{hji}(t)^{\Delta N_{hji}(t)} \right) \exp \left(\int_{0}^{c_i} \lambda_{hji}(u) du \right) \tag{15}$$

whereby, N_{hji} is the representation of a multivariate counting process for number of direct transitions between state h and state j for subject i within time $0 \longrightarrow t$, assuming that, unless in the presence of right-censored, the transition times are observed such that:

$$X(t), 0 \leq t \leq C_i \, ; i = 1 \dots \dots . n \tag{16}$$

$$N_{hji}, h, j \epsilon S, h \neq j, t \leq C_i \tag{17}$$

3.6.2. Choice-based models

In this study, three categories of choice-based models, binary, multinomial, and ordered models, are applied to accommodate different purposes. For all three categories, logistic and probit options are available; however, due to the underlying latent assumptions present in probit regression, and the straightforwardness in parameter interpretation, in terms of odds ratio for logistic regression, the literature favors the use of logistic (Woodridge 2012).

The binary based models consist of Logistic Regressions (LRs) and Rare Events Logistic Regressions (RELRs). The LRs are applied to model the pedestrians' use of pushbuttons, pedestrians' spatial crossing compliance, pedestrians' temporal crossing compliance, and drivers' spatial yielding compliance. In addition to the Traditional L, the partial-yield occurrence and occurrence of near-crash events are modeled using RELRs, then the results are compared. This is due to the rarity of the incidents, which raises a question of bias in coefficient estimation. However, there is still a debate on how rare the events must be in order to affect the coefficient estimates (Leitgöb 2013; Williams 2018).

3.6.2.1. Logistic Regression (LR)

In logistic regression, a dependent variable $\mathrm{Y}_i(i = 1, \ldots .n)$ follows a Bernoulli probability function with a value of 1 for probability θ_i and 0 for probability $1 - \theta_i$. The probability θ_i can be expressed as an inverse logistic function of a vector X_i of explanatory variables as:

$$\theta_i = \frac{1}{1 + e^{X_i \beta}} \tag{18}$$

The logistic function can be linearized and rewritten as shown in equation 19, whereby the $\hat{\beta}_s$ are the variables coefficients to be estimated including $\hat{\beta}_0$, which is a constant term:

$$logit(\theta_i) = ln\left(\frac{\theta_i}{1 - \theta_i}\right) = \hat{\beta}_0 + \hat{\beta}_1 X_1 + \ldots\ldots\ldots\ldots . + \hat{\beta}_n X_n \tag{19}$$

3.6.2.2.Rare Events Logistic Regression (RELR)

Although logistic regression has been extensively applied in modeling data with binary responses, it may result into extremely biased estimated coefficients when there is an imbalance of the proportion of response variables. This includes cases where the observed events of interest are very rare. Typical examples in this study are the occurrence of near-crash events and partial-yield states. In these cases, LR tends to underestimate the probability of event occurrence, as reported by King and Zeng (King and Zeng 2001). Thus, the rare event logistic regression is applied for bias correction (Guns and Vanacker 2012; King and Zeng 2001; Veazey et al. 2016). Basically, three steps are performed to modify LR to RELR (King and Zeng 2001).

i. The first is the resampling technique, in which all events (1s) are included in the sample and no events (0s) are selected randomly, in order to make a proportion of events (1s) to no events (0s) to be one to ten.

ii. Due to bias in the intercept term ($\widehat{\beta_0}$) , which is introduced in the first step, corrections that consider the fractions of events in population (ϕ) and in the sample (ψ) are taken into consideration:

$$\breve{\beta_0} = \widehat{\beta_0} - ln\left[\left(\frac{1-\phi}{\phi}\right)\left(\frac{\psi}{1-\psi}\right)\right] \tag{20}$$

iii. Lastly, the modifications that aim at correcting the underestimation of the probabilities are applied. This approach involves adding a correction factor C_i to the estimated probability $\widehat{p_\imath}$, which results into a new estimated probability (Equation 21). The correction factor C_i is computed as shown in equation 22, where X represents a 1×(n+1) vector of values for each independent variable β_i, X' is the transpose of X, and V($\breve{\beta_\imath}$) is the variance covariance matrix:

$$\breve{p_\imath} = \widehat{p_\imath} + C_i \tag{21}$$

$$C_i = (0.5 - \widehat{p_\imath})\widehat{p_\imath}(1 - \widehat{p_\imath})XV\left(\breve{\beta_\imath}\right)X' \tag{22}$$

In modelling spatiotemporal crossing compliance three types of models, Logistic Regressions (LRs), the Multinomial Logit (MNL) and Ordered Logistic Regressions (Ologit), are proposed. This is due to considerations in pedestrian crossing behaviors at signalized crosswalks with PHBs and TCSs. At these signalized crosswalks, pedestrian crossing behaviors can be categorized into four main groups, which are crossing within or outside the stripes, and crossing during WALK or DO NOT WALK signals. These behaviors can be grouped into three crossing compliance categories: full compliance, partial compliance, and non-compliance (Figure 19). Partial compliance includes pedestrians who complied with either the spatial or temporal requirements. Further, partial compliance and non-compliance can be grouped to form non-compliance when only pedestrians who crossed within the stripes when the WALK signal was active are considered compliant, as shown in Figure 19.

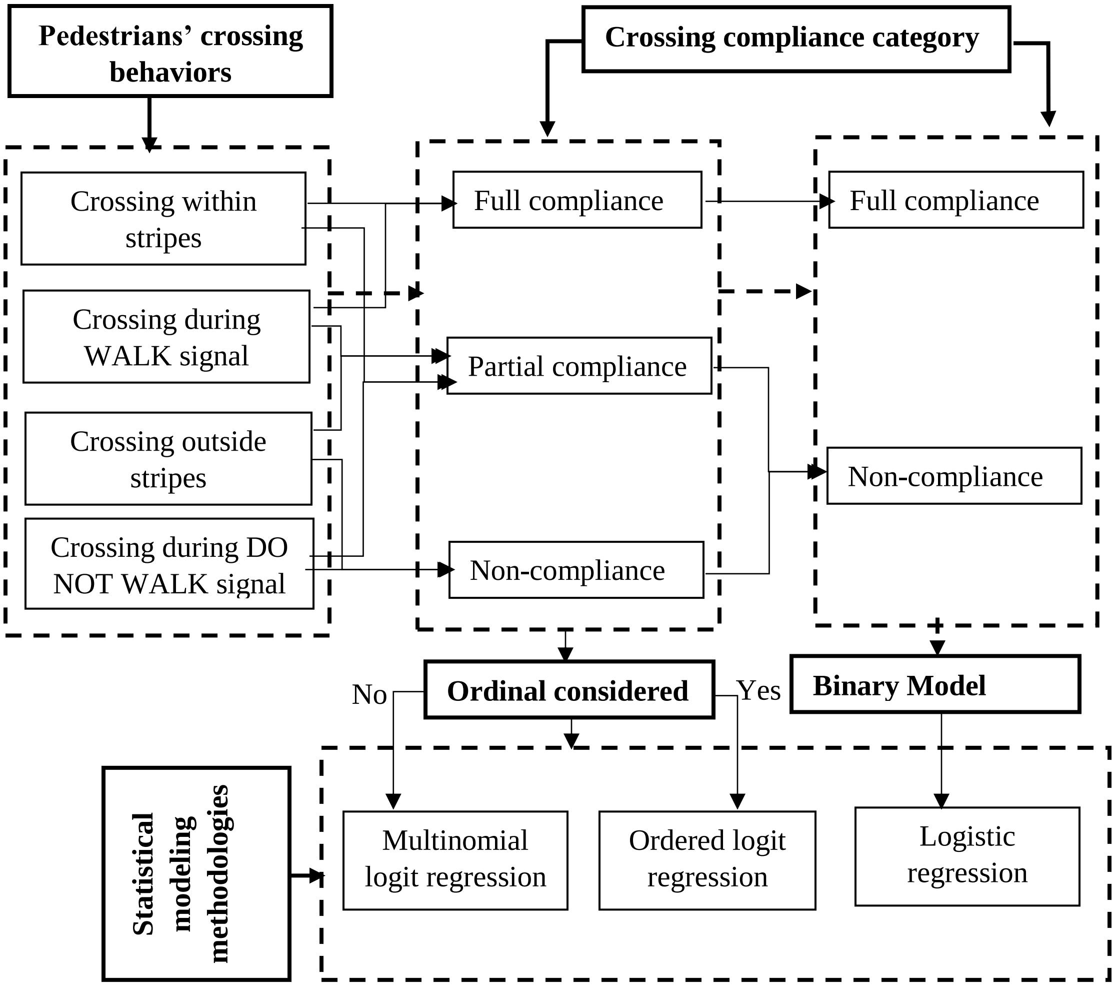


Figure 19. Models for spatiotemporal crossing compliance

The decision between using ordered and multinomial models depends on the consideration of ordinal scenarios for three compliance categories: full compliance, partial compliance, and non-compliance. If an ordinal scenario is considered, full compliance is considered at the highest rank, followed by partial compliance, and non-compliance becomes the lowest level. For this case, ordinal models can be used; otherwise, multinomial models are the best options. On the other hand, if partial compliance and non-compliance are grouped together, only two categories, full compliance and non-compliance, are formed, which calls for the application of binary-based models.

3.6.2.3. Multinomial Logistic Regression (MNL)

As described earlier, one approach to model spatiotemporal crossing compliance is Multinomial Logit (MNL). The MNL is applied when the outcome variable has more than two unordered categories. For this model, one of the categories is selected to be a base category. The probability of membership in any of the categories is compared to the probability of membership in the base category. For the outcome with M categories (three for this case), a total of M-1(two) equations are computed. The equation for each category of outcome variable in relation to the explanatory variables can be written as:

$$ln\frac{P(Y_i=m)}{P(Y_i=n)} = \hat{\beta}_0 + \sum_{k=1}^{k} \hat{\beta}_{mk} X_{ik} = Z_{mi} \tag{23}$$

Where X is the vector of variables, $\hat{\beta}$ is the vector of parameters to be estimated, n is the base category, and m is a non-base category.

In turn, the probabilities for the non-base category are computed as:

$$P(Y_i = m) = \frac{exp(Z_{mi})}{1 + \sum_{h=2}^{M} exp(Z_{hi})} \tag{24}$$

For the base category, the probabilities are computed as the reciprocals of the exponianted of each M-1 log odds:

$$P(Y_i = n) = \frac{1}{1 + \sum_{h=2}^{M} exp(Z_{hi})} \tag{25}$$

The key advantage of MNL models over ordinal models is that they do not impose unrealistic parameter restrictions. On the other hand, the vulnerability to the correlation of unobserved effects from one level to the next is the downside of the MNL models (Washington, Karlaftis, and Mannering 2011).

3.6.2.4. Ordered Logistic Regression (OLR)

The other approach to model spatiotemporal crossing compliance is by assuming that there is ordering/ranking in compliance. In this case, full complince is considered at the highest rank, followed by partial compliance, and non-compliance becomes the lowest level. Accounting for such an ordinal nature of the outcome variable, the ordered logit model is applied. The model derivation starts by specifying a latent variable Z, which is assumed to be a linear function of each spatiotemporal crossing observation:

$$Z = \hat{\beta} X + \varepsilon \tag{26}$$

where X is the vector of variables, $\hat{\beta}$ is the vector of parameters to be estimated, and ε is the error term. For ordered logistic regression, the error term is assumed to be logistically distributed (Washington, Karlaftis, and Mannering 2011). Thus, with the three ranks in the observed ordinal spatiotemporal crossing compliance data, for each observation, the y can be defined as:

$$\begin{array}{ll} y = 1 & if\ Z \le \mu_0 \\ y = 2 & if \mu_0 < Z \le \mu_1 \\ y = 3 & if \mu_1 < Z \le \mu_2 \\ y = \cdots \ldots .. & \\ y = 2 & if\ Z \ge \mu_{i-1} \end{array} \tag{27}$$

where μ are threshold parameters corresponding to the number of ranks, and are estimated jointly with model parameters $\hat{\beta}$.

3.6.3. Models' interpretation

3.6.3.1.Multistate model

The multistate models are interpreted by considering the magnitude and sign of the coefficient as well as the significance level. For simplicity, the coefficients are converted to hazard ratios, which are defined as the ratio of risk of outcome in the intervention group over the risk of outcome in the control group at a given interval of time, assuming that the subject in study has survived for a certain time. The hazard ratios are computed by exponentiation of the coefficients. The variables with positive coefficients (which in turn tend to have hazard ratios greater than one) are associated with a reduction of time to event, while negative coefficients (which in turn tend to have hazard ratios less than one) are associated with increased time to occurrence of an event. It should be noted that the hazard ratios provide the comparison of the time to event between two groups/variable categories, not the exact time elapsed for an event to occur (Spruance et al. 2004). In this study, the events of interest are the transitions of the traffic flow from non-yield state to full yield state, non-yield state to partial-yield state, and partial-yield state to full-yield state. Covariates that produce hazard ratios greater than one are desirable due to shortened duration, while those with hazard ratios less than one are undesirable.

3.6.3.2.Choice-based models

As the variable coefficients for choice-based models do not provide straightforward meanings, the model interpretation is based on the odds ratios, which are the exponents of the coefficients. Given regression results, if a variable has an odds ratio of greater than one, it implies that the presence of that variable is associated with the increased probability of occurrence of an event of interest. An odds ratio of less than one is associated with the decreased probability of an event of interest occurring, while an odds ratio of one can be interpreted as the variable has no significant impact on the event of interest (UCLA: Statistical Consulting Group 2014).

3.6.4. Model performance comparison

In model quality evaluation, for the information-based goodness of fit criteria, the Akaike Information Criterion (AIC) and Bayesian Information Criterion (AIC) are used. The AIC (Akaike 1974) is derived from information theory and chooses the model whose probability distribution discrepancy from the true distribution is the smallest. On the other hand, the BIC measures the trade-off between model fit and complexity (Stone 1979). The equations for AIC and BIC are given as (Fabozzi et al. 2014):

$$AIC = -2logL\left(\hat{\theta}\right) + 2k \tag{28}$$

$$BIC = -2logL\left(\hat{\theta}\right) + klog(n) \tag{29}$$

whereby θ stands for the vector of model parameters, $L\left(\hat{\theta}\right)$ is the likelihood of the model given data, k is the number parameters estimated by the model (slope, constant, variance etc), and n is the number of observations.

Both criteria are penalized by the addition of the new variable (parameter) into the model; however, the penalty is much higher for BIC than AIC. The model with a lower AIC or BIC value is prefered, since low BIC and AIC indicate a better fit (Kidando et al. 2017; Wang and Liu 2006). To decide on the best model, AIC or BIC scores between the competing models should be at least 10 (Fabozzi et al. 2014).

Since the comparison involves three discrete-based models, the author decided to further evaluate models in terms of their classification performances. A machine learning cross-validation criteria known as misclassification error rate was also included for selection of the best model. The misclassification error rate is the percentage of incorrect classified instances, given as:

$$e = \frac{FP + FN}{TP + TN + FP + FN} * 100 \tag{30}$$

whereby the FP = false positive, FN = false negative, TP = true positive , TN = true negative, and $e =$ misclassification error rate (%). The false positive and false negative are the incrorrectly classified incidents, while true positive and true negative are the correctly classified incidents. For the determination of the misclassification error, the dataset was divided into two groups: 60% of the data became the training set and was used for developing the model, while the 40% of the dataset was used to cross-validate the developed model.

3.6.5. Software for statistical modeling

In this study, the variables' coefficients were estimated by using the maximum likelihood method, as it is easily implementable in the available statistical software. The statistical modeling was performed in R version 3.5.1 environment (R Core Team 2018). Different packages were used, per different purposes. The MASS package (Ripley et al. 2018) and nnet package (Ripley and Venables 2016) were used for MNL, Ologit, and logistic model development. On the other hand, the p3state.msm package (Luis Meira-Machado et al. 2015) was used for developing the multistate hazard-based models. The "Zelig" package (Choirat et al. 2018) was used for RELR development, while the caret package (Kuhn 2017) was used for cross validation. Other packages used in analysis include "dplyr" (Wickham, François, et al. 2018) for data manipulations, and "ggplot2" (Wickham, Chang, et al. 2018) for plotting different graphs.

CHAPTER 4: DESCRIPTIVE ANALYSIS

A descriptive analysis of the data is purposely performed to provide a summary of the collected data. The descriptive statistics only make statements about the collected data, not the entire population. The summarized data is presented in terms of tables, charts, graphs, and figures. In associating two variables, the Pearson chi-square, which is a measure of association, was used to evaluate the relationship between them, with the level of significance (α) set at 0.05. The descriptive analysis chapter is divided into four main parts: general descriptive statistics; yielding compliance of drivers; pedestrian compliance to crosswalk features; and descriptive summary of the potential variables.

4.1.General descriptive statistics

This section covers the general descriptive analysis of the data including the distribution of the observations per data collection site as well as per signal type.

Table 2. Number of observations per site

	Site location		**Signal Type**	**Number of observations**
sn	**Main street**	**Closest minor street**		
1	Boulder	Sun Valley	RRFB	117
2	Charleston	11th Street	RRFB	149
3	Charleston	17th Street	RRFB	113
4	Charleston	Lamont	RRFB	106
5	Commerce	La Madre	RRFB	118
6	Craig	Ferrell	RRFB	105
7	Flamingo	Cameron	RRFB	109
8	Flamingo	Linq Ln	TCS	104
9	Flamingo	Mojave	RRFB	105
10	Las Vegas Blvd	Convention Center	TCS	133
11	Las Vegas Blvd	Welcome sign	TCS	138
12	Maryland Pkwy	Del Mar st	TCS	114
13	Maryland Pkwy	Dumont Blvd	CFB	138
14	Maryland Pkwy	University Ave	CFB	198
15	Maryland Pkwy	Reno Ave	CFB	134
16	Sahara	15th Street	PHB	161
17	Sahara	Las Verdes	RRFB	128
18	Swenson	South Dr	CFB	149
19	Valley view Blvd	Conlon Ave	CRFB	109
20	Warm Springs Rd	Giles Street	CFB	210

As described earlier, the number of observations for each site varied between 100 and 300. The site with the largest observations was located on Warm Springs road near Giles street. On the other hand, the site with the fewest observations was on Flaming near Linq Ln. The number of observations for each site are as shown in Table 2.

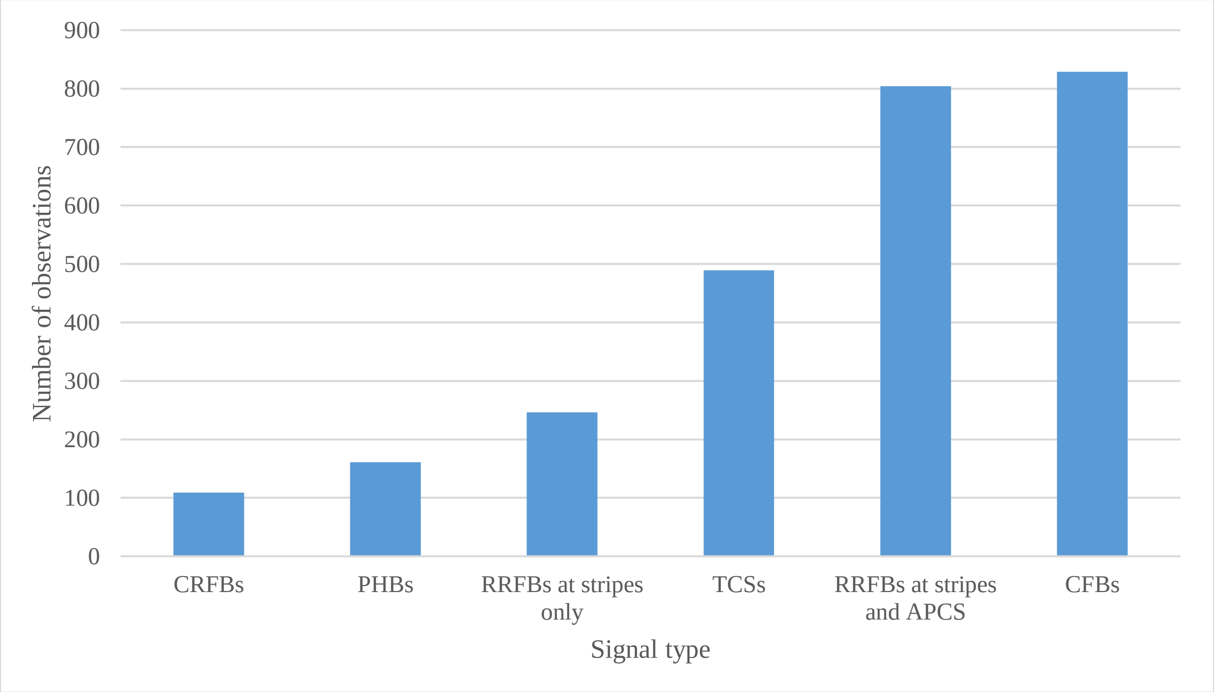


Figure 20. Number of observations per signal type

In general, a total of 2638 observations were recorded for all 20 sites, including four CFBs, nine RRFBs, one CRFB, four TCSs, and one PHB. From the total observations, 829 observations were recorded from sites with CFBs, 1050 observations from RRFB sites (with different signal arrangements), 161 from PHB sites, and 109 from CRFB sites, while 489 observations were from TCS sites (Figure 20). Most of the observations were collected from CFBs and RRFBs, since they are the common signal types in the study area. The least number of observations were from CRFBs and PHBs, as each of these signal types had only one site.

4.2. Yielding compliance of drivers

The yielding compliance of drivers is assessed based on whether partial-yield and full-yield occurred, the yielding zone where the full-yield occurred, and the time taken to yield. These yielding criteria are assessed with respect to the direction of travel and signal types, as well as the presence of supplementary markings and signs at the crosswalks.

4.2.1. Partial-yield and full yielding incidents

As defined earlier, partial-yield events involve incidents in which a portion of the drivers stop, while others continue driving. Conversely, full-yield incidents are those in which all drivers yield right-of-way to the pedestrians who are waiting to use the crosswalk. Yielding right-of-way, for this case, is considered as stopping or reducing vehicle speed, after either the traffic signal changes to red or the flashing lights begin. In the entire dataset, about 32% of the pedestrians crossing incidents did not involve drivers yielding right of way. In Figure 21, both partial-yield events and full-yield incidents are presented according to the signal types at which they occurred. It can be observed that TCSs, with 89.2%, and PHBs, with 80.4%, are the signal types that have the highest full-yield compliance. The same signal types have the lowest partial-yield frequencies. Considering flash-based signals, the CFBs have the lowest full-yield rates (58.5%), while CRFBs have the highest yield rate (78.9%). The highest percentage of partial-yield (7.3%) is observed at CRFBs, while the lowest (0%) is at TCSs. Another observation from Figure 21 is that the flash-based signal types have relatively low full-yielding incidents, but a high number of partial-yield incidents, as compared to TCS and PHB signals.

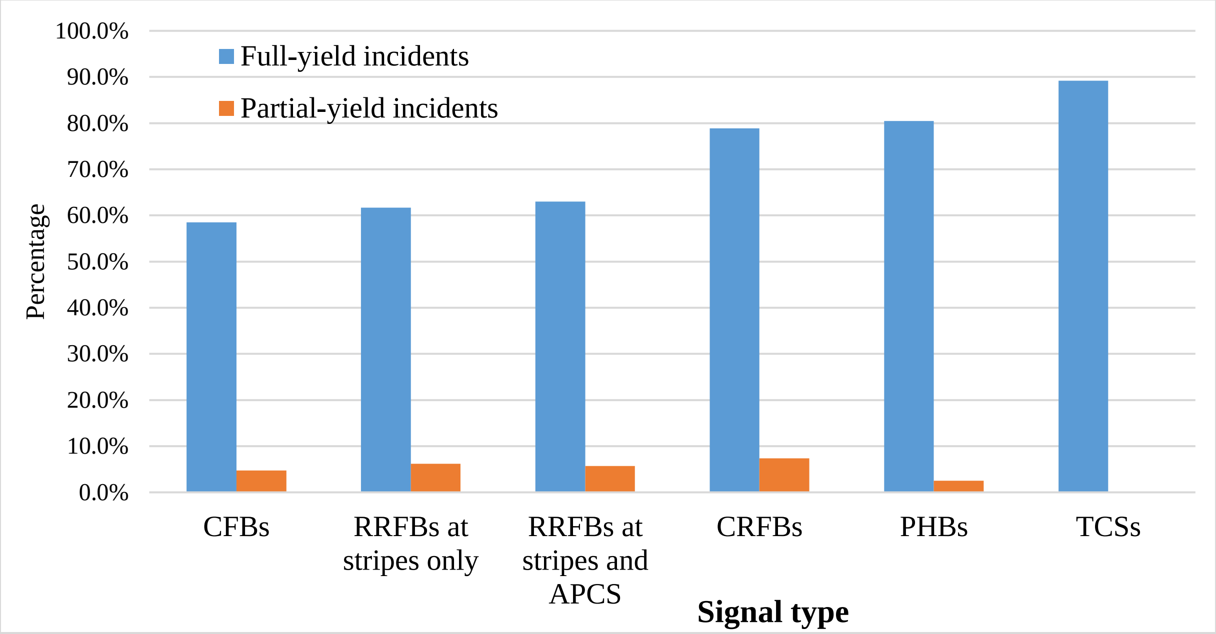


Figure 21. Partial-yield and full yielding incidents across signal types

4.2.2. States' transition durations

The time taken for state transition during pedestrian-driver interaction at a crosswalk is an important ingredient in pedestrian safety. It is assumed that the longer it takes for a non-yield to full-yield compliance transition to occur, the higher the possibility that a pedestrian will jaywalk. Therefore, the better performing signal is not only the one that has a large proportion of vehicles yielding to pedestrians, but also reduces the time to yield. The descriptive summary for state transition durations, which includes the transitions from non-yield states to full-yield states, non-yield states to partial-yield states, and partial-yield states to full-yield states, for various crosswalk signals, is presented in Table 3. Further analyses considering the distribution of transition durations for various signal types are presented in Figures 22 through 24.

4.2.2.1.Summary of transition durations

According to Table 3, it can be observed that, on average, drivers take a short time to yield right-of-way for CFBs, CRFBs, and RRFBs, as compared to PHBs and TCSs. This may be according to the design, since PHBs and TCSs are synchronized to other traffic signals in the network. The comparison across flash-based signals shows the RRFBs have the longest average and maximum durations of traffic transitioning from non-yield to full-yield.

Table 3. Summarized transition times across signal types

Signal type	Change of state	Observations	Transition duration			
			Average	Std dev	Min	Max
CFBs	Non-yield to partial-yield state	79	12	13	1	81
	Partial-yield to full-yield state	79	5	2	1	13
	Non-yield to full-yield state	891	12	10	0	82
CRFBs	Non-yield to partial-yield state	16	14	4	7	19
	Partial-yield to full-yield state	16	6	2	3	8
	Non-yield to full-yield state	156	14	8	0	33
RRFBs at stripes only	Non-yield to partial-yield state	28	8	5	3	18
	Partial-yield to full-yield state	28	4	1	2	6
	Non-yield to full-yield state	282	17	18	0	119
RRFBs at stripes and APCS	Non-yield to partial-yield state	100	15	13	2	58
	Partial-yield to full-yield state	100	5	3	1	15
	Non-yield to full-yield state	892	15	17	0	262
PHBs	Non-yield to partial-yield state	8	55	31	9	105
	Partial-yield to full-yield state	8	10	12	3	38
	Non-yield to full-yield state	251	65	58	0	318
TCSs	Non-yield to partial-yield state	0	0	0	0	0
	Partial-yield to full-yield state	0	0	0	0	0
	Non-yield to full-yield state	872	57	50	0	277

The crosswalks with RRFBs at the stripes and APCS have the largest maximum time (262 seconds), while crosswalks with RRFBs at the stripes only have the largest average time (17 seconds) of transitioning. Considering the non-yield to partial-yield change of state, the maximum time was 81 seconds, which was observed at the crosswalk equipped with a CFB signal. The maximum average time was 15 seconds, which was observed at the crosswalk with RRFBs at the stripes and APCS.

The transition from a partial-yield to full-yield state took a relatively short period of time. On average, this transition took between four to five seconds for flash-based signals. For TCSs and PHBs, the longest time of transition from non-yield directly to full-yield was 318 seconds, while the average was 57 seconds and 65 seconds for TCSs and PHBs, respectively. No partial-yield state was observed for TCSs, while only 10 observations from PHBs were partial-yield states, which took about 48 seconds to occur.

4.2.2.2.Distribution of states' transition durations

The density plots presented in Figures 22 through 24, which show the distributions of the durations for different transitions, reveal that there is uniformity in the time of non-yield to full-yield transitions for flash-based signals. This can be observed in Figure 22, where the transition time graphs are closely packed for all the flash-based signals. The TCSs and PHBs, on the other hand, not only have significantly long durations for this transition, but also significantly different durations, as shown by the variability of their graphs. Each signal type exhibits different duration distributions when non-yield to partial-yield state transitions (Figure 23) and partial-yield to full-yield state transitions (Figure 24) are considered.

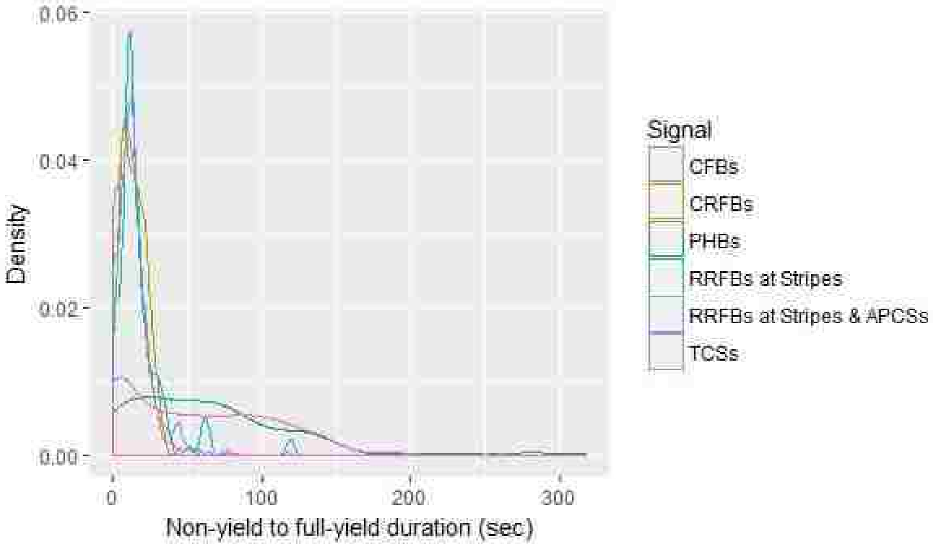


Figure 22. Non-yield to full-yield transition durations across signal types

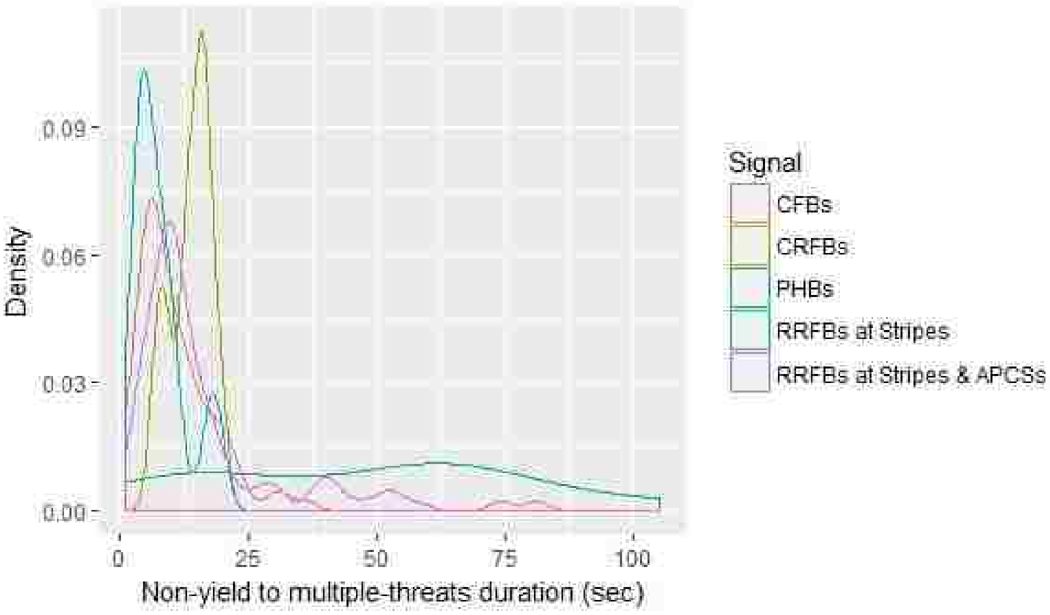


Figure 23. Non-yield to partial-yield state transition durations across signal types

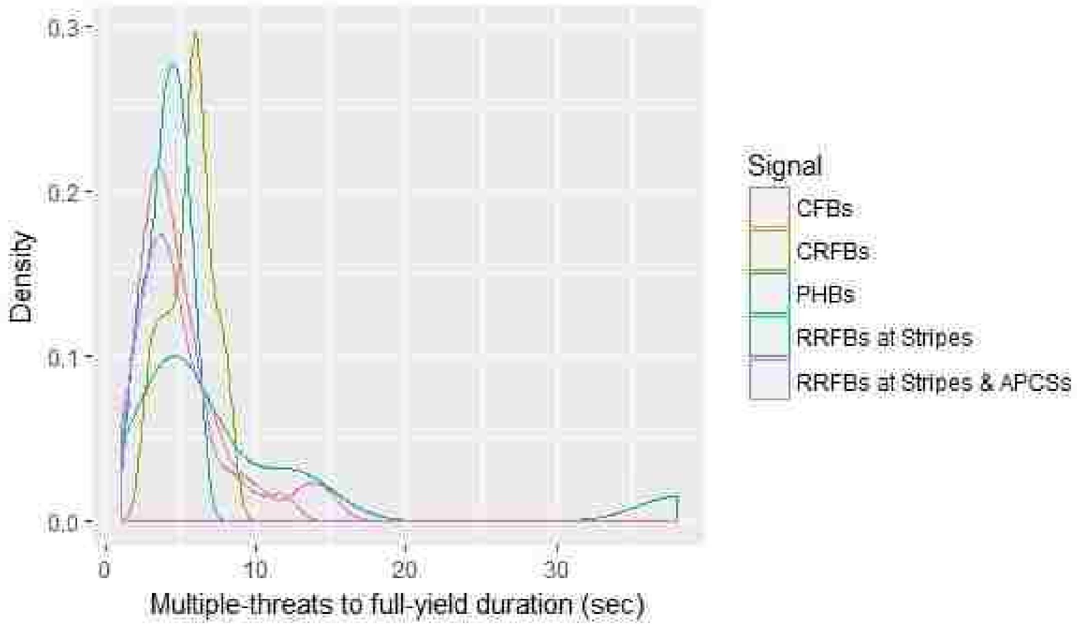


Figure 24. Partial-yield to full-yield state transition durations across signal types

4.3.Pedestrians' compliance to crosswalk features

In this section, the summary of pedestrian compliance towards the effective utilization of crosswalk features that are designed to enhance safe crossing of the roadways is presented. The main features are the pushbuttons, as well as "LOOK BEFORE CROSSING," "USE CROSSWALK," and "PUSH BUTTON TO TURN ON LIGHTS," sign boards. Pushbuttons are used to activate flashing lights for RRFB, CFB, and CRFB signals at the crosswalks, and to request a walk phase for TCS and PHB signals. A high pushbutton activation rate translates to compliance in pushbutton usage. This section attempts to investigate whether the presence of the signs that instruct pedestrians to press the pushbutton improve pushbutton activation rates. The greater the number of people utilizing the crosswalk, instead of jaywalking, translates into high compliance to the "USE CROSSWALK" signs. Similarly, the presence of a large proportion of pedestrians looking before crossing implies that they are complying with the "LOOK BEFORE CROSSING" directives provided. Therefore, the question that needs to be answered is, to what extent do the presence of signs that direct pedestrians to perform certain actions before crossing actually alter pedestrian behaviors. To determine this, statistical tests were performed to evaluate behavioral changes for crosswalks both with and without signs.

4.3.1. Pushbutton pressing compliance

Overall, 58.72% of pedestrians and bicyclists pushed the button, irrespective of their arrival sequence (Table 4). However, such a statistic may be misleading, due to the fact that the pedestrians who arrived while other pedestrians were either crossing or waiting to cross, could have assumed that the button had already been pressed; thus, there would be no need for them to repeat a similar action. When only pedestrians who were first to arrive at the crosswalk are considered, the overall pushbutton pressing rate increased to 70%. There exists a variation of pushbutton pressing rates across the signal types, with PHBs having the highest rate at 72.8%, while TCSs and CFBs have the lowest rates, both at 66%. An interesting observation is at the TCS signal type, when only the first arriving pedestrians are considered, the activation rate (66.41%) is almost 1.5 times that of when all pedestrians, irrespective of their arrival sequence, are considered (40.29%). This can be attributed to the presence of a significant number of first-arriving pedestrians and follow-up pedestrians. Between 10% and 32% of the follow-up pedestrians pressed the pushbutton before crossing at crosswalks with different signal types.

Table 4. Pushbutton activation rates

Pedestrian action	Overall	CFBs	CRFBs	RRFBs	PHBs	TCSs
	All pedestrians					
Pressed	58.72%	55.01%	90.83%	66.67%	60.25%	40.29%
Did not press	41.28%	44.99%	9.17%	33.33%	39.75%	59.71%
	Pedestrian who were the first to arrive at the crosswalk					
Pressed	70.00%	66.40%	94.12%	70.43%	72.80%	66.41%
Did not press	30.00%	33.60%	5.88%	29.57%	27.20%	33.59%
	Pedestrian who were not the first to arrive at the crosswalk					
Pressed	17.61%	17.95%	42.86%	32.04%	16.67%	10.13%
Did not press	82.39%	82.05%	57.14%	67.96%	83.33%	89.87%

Further analysis results presented in Table 4 show that among first arriving pedestrians at locations with “PUSH BUTTON TO TURN ON LIGHTS” signs, 71.4% pressed the pushbutton, while 55.6% did the same for locations without the signs, which makes a difference of about 16%.

Table 5. Pushbutton activation rates with and without signage

			“PUSH BUTTON TO TURN ON LIGHTS" present	
	Pressed pushbutton		**No**	**Yes**
First to arrive pedestrian(s)	No	Count	80	541
		Percentage	44.4%	28.6%
	Yes	Count	100	1,349
		Percentage	55.6%	71.4%
	Total	Count	180	1,890
		Percentage	100.0%	100.0%
		Pearson chi2(1) = 19.6	P-value = 0.000	
Follow up pedestrian(s)	No	Count	64	404
		Percentage	90.1%	80.3%
	Yes	Count	7	93
		Percentage	9.9%	19.7%
	Total	Count	71	497
		Percentage	100.0%	100.0%
		Pearson chi2(1) = 3.4	P-value = 0.067	

For the follow up pedestrians, 9.9% (7 out of 71 pedestrians) pressed a pushbutton in the absence of a sign, while 93 out of 404 (19.7%) performed a similar action in the presence of a sign. Table 5 shows that the Pearson chi square for the two variables is 19.6 for pedestrians who were first to arrive at the crosswalk, which is a very strong association at a 95% confidence level, by which the critical value for one degree of freedom is 3.84. Moreover, for pedestrians who were not the first to arrive at a crosswalk, the association is not statistically significant at a 95% level.

4.3.2. Look before crossing

Of the 20 crosswalks, only four were equipped with “LOOK BEFORE CROSSING” signs. Table 6 shows that almost all pedestrians looked in both directions of traffic flow before crossing, regardless of the presence of the signs. The difference in looking for oncoming traffic in the presence and absence of the signs was only about 0.4%. The Pearson chi-square results show that there is no statistically significant difference, at a 95% confidence level (P-value =0.075), in looking before crossing, irrespective of the presence or absence of the “LOOK BEFORE CROSSING” sign. This finding suggests that the presence of a “LOOK BEFORE CROSSING” sign does not alter pedestrian behavior towards looking for oncoming traffic.

Table 6. Pedestrians looked before crossing

		“LOOK BEFORE CROSSING” sign present	
Looked before crossing		**No**	**Yes**
No	Count	9	0
	Percentage	0.4%	0.0%
Yes	Count	1,945	684
	Percentage	99.6%	100.0%
Total	Count	1,954	684
	Percentage	100.0%	100.0%

Pearson chi2(1) = 3.1612 P-value = 0.075, 1-sided Fisher's exact = 0.067

4.3.3. Pedestrians’ spatial crossing compliance

The pedestrian spatial crossing compliance rate (SCCR) is the rate at which pedestrians comply to use a designated path within a crosswalk. According to most state laws, including Nevada, drivers are supposed to yield to pedestrians who are already in the crosswalk; however, the definition of crosswalk is not clear. For instance, it is not clear whether the defined “crosswalk” means the striped locations, the distance between the yield lines, or the distance between the advanced pedestrians crossing signs. In this study, crossing compliance implies crossing within the striped marks for both directions of traffic flows.

Table 7. Pedestrians spatial crossing compliance rates (SCCRs)

			Destination			
	Crossing zone		Within marked stripes	Between stripes and yield line	Between yield line and APCS	Total
Origin	Within marked stripes	Count	2120	113	7	2240
		Percentage	80.36%	4.28%	0.27%	84.91%
	Between stripes and yield line	Count	113	43	11	167
		Percentage	4.28%	1.63%	0.42%	6.33%
	Between yield line and APCS	Count	42	9	180	231
		Percentage	1.59%	0.34%	6.82%	8.76%
	Total	Count	2275	165	198	2,638
		Percentage	86.24%	6.25%	7.51%	100.00%

The findings presented in (Table 7) show that about 80% of pedestrians fully complied to cross within the stripes, while about 10.42% partially complied. The partially complying pedestrians include: about 4.28% who started crossing within marked stripes, but ended their crossings in the zone that is between stripes and yield line; 4.28% who started crossing between the stripes and yield line, but finalized crossing within marked stripes; 0.27% who started crossing within the marked stripes and finished between the yield line and APCS; and 1.59% who started between the yield line and APCS and finalized within the marked stripes. Contrarily, about 9.22% of pedestrians did not comply with crossing within designated locations, as they started and finished their crossings in zones that were outside of the marked stripes.

To further explore the influence of the “USE CROSSWALK” signs on pedestrian crossing compliance, an association analysis was performed. Since the “USE CROSSWALK” signs are normally placed on the sides of roadways, only the side where the pedestrian originated was considered. The results in Table 8 show a very large percentage of pedestrians who crossed within the stripes in both the presence (87.4%) and absence (84.7%) of the sign. This is an indication that the presence of the sign at the crosswalk is not statistically significant when associated with pedestrians’ spatial crossing compliance.

Table 8. Influence of “USE CROSSWALK” sign on crossing compliance

			“USE CROSSWALK” sign present	
			No	**Yes**
Crossing zone	Within stripes	Count	2,045	195
		Percentage	84.7%	87.4%
	Outside stripes	Count	370	28
		Percentage	15.3%	12.6%
	Total	Count	2,415	223
		Percentage	100.0%	100.0%

Pearson chi2(2) = 1.3 P-value = 0.524

This is further revealed by the Pearson chi square result of 1.3, which is below critical value (3.84), at a 95% level. Therefore, it can be concluded that no statistically significant difference in results for pedestrians crossing behaviors were observed in either the presence or absence of "USE CROSSWALK" signs.

4.3.4. Pedestrians temporal crossing compliance

This part of the study used only 650 crossing incidents that occurred at TCSs and PHBs, since at these signal types pedestrians are supposed to wait for their crossing phase. The temporal crossing compliance rate for this case is defined as the percentage of crossing incidents in which pedestrians waited for their crossing phase (Sisiopiku and Akin 2003). The general statistics (Table 9) show that 71.2% of crossing incidents involved pedestrians who complied to wait for the crossing phase in the presence of a pushbutton with a "WAIT" voice, while in the absence of the device, 70.5% waited for the crossing phase. However, since the audible devices can only be heard if a person presses the pushbutton, it is logical to consider only the pedestrians who pressed the pushbutton for analysis.

Table 9. General temporal crossing compliance statistics

Temporal comply		**Audible device present**	
		No	**Yes**
No	Count	70	119
	Percent	29.5%	28.8%
Yes	Count	167	294
	Percent	70.5%	71.2%
Total	Count	237	413
	Percent	100.0%	100.0%

Table 10 below presents the influence of audible devices on temporal crossing compliance rates for pedestrians who pressed the pushbutton. It can be observed that although there seems to be a high percentage of pedestrians who complied to wait after pushing the pushbutton (73.0%) in the presence of an audible device, a high percentage (67.1%) of temporally complying pedestrians is also observed in the absence of the device. Therefore, the association between the presence of audible devices and temporal crossing compliance is very weak. This is further revealed by the Pearson chi-square association coefficient of 0.997, which is statistically at a 95% confidence level. Therefore, it can be concluded that the audible devices have no statistically significant influence on temporal crossing compliance.

Table 10. Temporal crossing compliance rates with respect to audible devices

Temporal comply		**Audible device present**	
		No	**Yes**
No	Count	26	58
	Percent	32.9%	27.0%
Yes	Count	53	157
	Percent	67.1%	73.0%
Total	Count	79	215
	Percent	100.0%	100.0%

Pearson chi2(1) = 0.9970 P-value = 0.318

4.3.5. Pedestrians spatiotemporal crossing compliance

The variation of spatiotemporal crossing compliance across the signal types was further investigated. According to Figure 25, TCSs are observed to have a higher spatiotemporal crossing compliance than PHBs. About 67% of pedestrian crossing incidents at TCSs complied to both spatial and temporal crossing rules, while only 52% did the same for PHBs. On the other hand, about 43% of crossing incidents were considered complient to either spatial or temporal rules for PHBs, compared to 28% for TCSs. For the total amount of incompliant incidents (neither spatial nor temporal), both signal types performed relatively similarly, as shown by 5% for PHBs and 4.5% for TCSs.

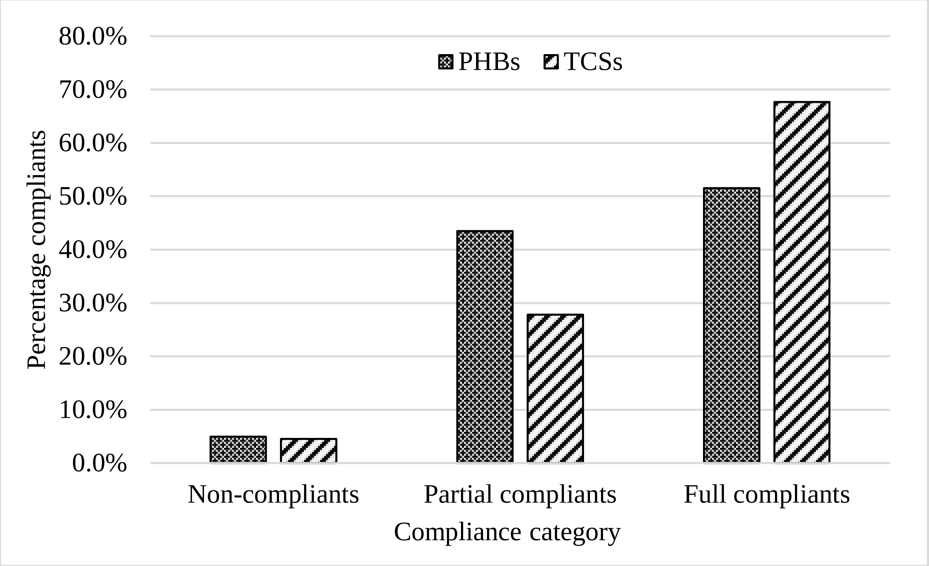


Figure 25. Spatial and temporal crossing compliance across signal types

4.4.Descriptive summary of potential variables

The variables presented in this section can be categorized as dependent and explanatory variables, which are then subcategorized into continuous, binary, and categorical variables. The descriptive analysis of these variables shows the number of observations for each variable, and the percentage composition of observations for binary and categorical variables. Meanwhile, for continuous variables, the descriptive summary includes the average and standard deviation, as well as minimum and maximum values. The potential explanatory variables are subdivided into crosswalk-related, pedestrian-related, traffic-related, and temporal-related variables, as presented below.

4.4.1. Count/Temporal based dependent variables

The count/temporal based variables include both the time and vehicles that passed, from when the pedestrians arrived at the crosswalk to the moment that all of the vehicles stopped for pedestrians. Table 11 shows the variations of these variables. On average, the time elapsed from non-yield to full-yield was about 29 seconds if no partial-yield occurred, and 20 seconds if partial-yield occurred. On the other hand, the average time from non-yield to partial-yield was 15 seconds, and partial-yield to full-yield was 5 seconds. The maximum time from non-yield to full-yield was 318 seconds, non-yield to partial-yield was 105 seconds, and partial-yield to full-yield was 38 seconds. There were 1701 incidents in which driver yielding was not observed. In those incidents, pedestrians waited for an average of 11 seconds, with a maximum wait time of 397 seconds. Furthermore, during those state transitions, the number of vehicles that passed were counted and are presented in Table 11. On average, the third vehicle stopped for pedestrians to cross. The maximum number for non-yield to full-yield transition was 55, while that of non-yield to partial-yield was 12. In the incidents in which no yielding was observed, a maximum of 26 vehicles were observed before the pedestrians used the crosswalks.

Table 11. Continuous dependent variables

Variable	Description	Obs	Mean	Stdev	Min	Max
Non-yield to full-yield time	Time until full-yield occurred	3344	29	39	0	318
Non-yield to partial-yield time	Time until partial-yield occurred	231	15	15	1	105
Partial-yield to full-yield time	Time from partial-yield to full-yield	231	5	4	1	38
Non to partial-yield to full-yield	Time from non-yield to full-yield through partial-yield	231	20	16	2	117
Full-yield time	Time until full-yield occurred, no partial-yield considerations	3575	28	38	0	318
No yield time	Time elapsed where no yield occurred	1701	11	39	0	397
Full yield vehicles	Number of vehicles passes before full yield	3575	2	5	0	55
Partial-yield vehicles	Number of vehicles passes during partial-yield	231	2	1	1	12
No yield vehicles	Number of vehicles passed where no yield occurred	1701	0.1	1.2	0	26

4.4.2. Categorical-based dependent variables

The binary-based dependent variables show whether a certain action happened. Some variables are pedestrian-based, while others are vehicular traffic-based. In light of the pedestrian-based variables, according to Table 12, on average about 59% of the pedestrians pushed the button before crossing. This rate was computed irrespective of the arrival sequence of the pedestrians or the status of the flashing lights. The temporal crossing compliance shows whether pedestrians waited for the “WALK” signal at PHB and TCS signalized crosswalks. In this study, about 82% of pedestrians waited for the signal. Parallel to that, the spatial crossing compliance is described as the percentage of pedestrians that crossed within the marked stripes, which is considered as the dedicated path during crossing. The spatial crossing compliance for this study was around 85%.

The traffic-based variables include the full-yield, partial-yield, near-crash, and yield compliance. The full-yield variable represents the stopping or reducing of vehicle speeds for pedestrians to use the crosswalks. On average, in 67.76% of the incidents, traffic flow stopped for pedestrians. Partial-yield, which describes an incident in which a portion of the vehicles in the traffic flow stopped while others did not, constituted of 4.55% of all incidents, while a total of 66 (1.25%) near-crash events occurred.

The spatial yielding compliance, which represents a situation in which traffic stopped before the yield line, accounted for 78.29%. As binary, multinomial, and ordered models are developed for modeling spatiotemporal yielding compliance, the corresponding variables need more clarification. There were few observations of this variable, since the spatiotemporal yielding compliance applies only for TCSs and PHBs. Under multinomial and ordered models, it can be observed in Table 12 that pedestrians in 4.62% of crossing incidents were non-compliant (neither complied spatially nor temporally). Moreover, 31.69% of crossing incidents involved partially-compliant pedestrians (complying either spatially or temporally), while 63.69% of the incidents involved fully-compliant pedestrians. On the other hand, under binary models for spatiotemporal crossing compliance, partially compliant and non-compliant pedestrians are grouped together to form the non-compliant group, whose percentage then become 36.31%.

Table 12. Categorical dependent variables

Variable	Description	Code and category	Count	Percent
Pressed	Whether pedestrian pressed pushbutton	(1) Yes (0) No	1549	58.72%
Temporal compliance	Whether pedestrians waited for walk signal	(1) Yes (0) No	429	81.97%
Spatial compliance	Whether pedestrians crossed within marked stripes	(1) Yes (0) No	1120	84.91%
Full-yield	Whether full-yield occurred	(1) Yes (0) No	3575	67.76%
Spatial yield compliance	Whether vehicles yielded before yield line	(1) Yes (0) No	3113	78.29%
Partial-yield	Whether partial-yield occurred	(1) Yes (0) No	231	4.55%
Near crash	Whether a near-crash event was observed	(1) Yes (0) No	66	1.25%
Spatiotemporal Crossing compliance (MNL&Ologit)	Whether pedestrians complied to spatiotemporal crossing	(1) Non-compliants	30	4.62%
		(2) Partial compliants	206	31.69%
		(3) Full compliants	414	63.69%
Spatiotemporal Crossing compliance (Logistic regression)	Whether pedestrians complied to spatiotemporal crossing	(0) Non-compliants	236	36.31%
		(1) Full compliants	414	63.69%

4.4.3. Crosswalk-related explanatory variables

Table 13 presents the descriptive analysis of the crosswalk-related variables. Considering signal types, RRFBs, and CFBs have relatively large percentages of observations compared to other signal types, since most of the crosswalks had these signal types.

Table 13. Crosswalk related variables

	Variable type and name	Description	Count	Percent
Categorical variables	Signal type			
	CFBs	Type of signal installed	828	31.39%
	CRFBs		109	4.13%
	PHBs		162	6.14%
	TCSs		489	18.54%
	RRFBs at stripes only		246	9.33%
	RRFBs at stripes and APCS		804	30.48%
	Cross stages			
	Two stages	The crossing stages at the crosswalk	404	15.31%
	Optional one/two		755	28.62%
	Strictly one		1479	56.07%
	Crash history			
	Less than 10 crashes	Number of crashes per year (2014-2016)	854	32.37%
	Between 10 and 20 crashes		1072	40.64%
	More than 20 crashes		712	26.99%
	Number of lanes			
	Five	Number of lanes at the crosswalk	401	15.20%
	Six lanes		771	29.23%
	Seven lanes		627	23.77%
	Eight to ten lanes		839	31.80%
	Land use			
	Mixed	Land use where crosswalk is located	1064	40.33%
	Residential		840	31.84%
	Commercial		734	27.82%
	Yield line to marked stripes distance			
	Less than 40ft	Yield line to marked stripes distance	1942	36.81%
	Between 40 and 80ft		2266	42.95%
	More than 80ft		1068	20.24%
	Median type			
	No or TWTL	Median type	615	23.31%
	Narrow raised		1474	55.88%
	Wide raised		549	20.81%
	Yield line to APCS distance			
	Less than 100ft	Distance from yield line to Advanced pedestrian crossing sign	1942	36.81%
	Between 100 and 200ft		2266	42.95%
	More than 200ft		1068	20.24%
Binary variables	State's law sign	Whether "STATE'S LAW" sign was present	272	10.31%
	Ped Xing sign	Whether "PED XING" sign was present	214	8.11%
	Use crosswalk sign	Whether "USE CROSSWALK" sign was present	223	8.45%
	Yield here sign	Whether "YIELD HERE" sign was present	1932	73.24%
	Inside and outside turn lanes	Presence of inside and outside turning lanes	931	35.29%
	Activated/active flashes	Whether lights were flashing or red	1904	72.18%
	Green light	Whether next intersection lights were green	1390	26.35%

About 56% of the observations were recorded at crosswalks with strictly one cross stage, while the strictly two cross stages had about 15%. It should be noted that cross stages represent the number of stops that pedestrians have to undergo before they reach the second side of the road. Some crosswalks are deliberately designed with one cross stage, whereby no refuge is provided at the median; others have two stages by provision of refuge at the median; while in other designs the two-stage crossing is optional for pedestrians.

The crash history variable shows that most of the crosswalks are located where 10 to 20 crashes (40.64%) occurred between 2013 and 2016, followed by less than 10 crashes (32.37%), and lastly, more than 20 crashes (26.99%). The number of lanes for the crosswalks varied from five to ten, whereby the composition percentage of the observations varied from about 15% for five-lane crosswalks, to almost double (31%) for the locations with eight to ten lanes. Mixed land use constitutes the largest proportion of the observations in this study, followed by commercial and residential land uses. The narrow raised median type has a relatively large proportion of the observations, compared to wide-raised, and Two-Way Turn Lanes (TWTL).

The distance between the marked stripes and advanced pedestrian crossing signs (APCSs) was of interest. This distance was subdivided into two zones: the yield line to marked stripes distance, and the yield line to APCS distance. The 40ft distance was considered as a benchmark, since it is provided in the MUTCD as the minimum distance at which to locate the yield line. The distribution of observations for the yield line to the marked stripes distance is 36.81%, 42.95%, and 20.24% for less than 40 ft, between 40 ft and 80 ft, and more than 80 ft, respectively. On the other hand, for the yield line to APCS, the distance was categorized as less than 100 ft, between 100 and 200 ft, and more than 200 ft with 36.81%, 42.95%, and 20.24% of observations, respectively.

Considering the signs at the crosswalks, 10.31% of observations were recorded at crosswalks with “STATE'S LAW YIELD TO PEDESTRIANS” signs, 8.11% with “PED XING” signs, 8.45% with “USE CROSSWALK” signs, and 73.24% with “YIELD HERE TO PEDESTRIANS” signs. About 35% of the observations were from locations with inside and outside turning lanes. The flashing lights and nearby intersection signal status were also of interest, as they impact both pedestrians’ and drivers’ behaviors.

4.4.4. Pedestrian-related variables

Several pedestrian-related variables were recorded at the sites, and their descriptive summary is presented in Table 14. Considering gender, more than half (58.9%) of the crossing incidents involved males only, while nearly a quarter (24.3%) involved females only; the incidents where males and females crossed together accounted for 16.8% of all observations.

The approximate ages of pedestrians were estimated visually, whereby four groups were identified. If the crossing incident involved more than one age group, it was categorized as mixed ages. Among all of the age groups, adults accounted for the highest percentage (55.4%) of the crossing incidents, followed by the young adults’ group (19.4%), and children and teens (11.6%), with the elderly only group having only 4% of the crossing incidents. Most of the pedestrians (84.91%) started crossing within the marked stripes, and even more (86.24%) finalized within the same zone. On the other hand, 8.76% and 7.51% started and finished crossing, respectively, between the yield line and APCS.

Based on pedestrian activities before and during crossing, a relatively large percentage of pedestrians (77.79%) and (65.43%) were walking normally before and during crossing, respectively. There was an increase in the percentage of pedestrians who rode bikes during crossing (19.79%) compared to before crossing (7.66%). The percentage of phone use was relatively lower during crossing (1.18%) than before crossing (1.90%). Most of the crossing incidents (71%) involved one pedestrian crossing, while only about 7% involved three or more pedestrians. More than three quarters of the crossing incidents involved pedestrians who arrived at the crosswalk while no one else was using it. Only a few pedestrians were coming or going to a bus, and even fewer approached the crosswalk from the far side.

Table 14. Pedestrians' related variables

	Variable type and name		Description	Count	Percent
Categorical	Pedestrian crossing zone 1		Zone in which pedestrian start crossing		
		Within stripes		2240	84.91%
		Between stripes and yield line		167	6.33%
		Between yield line and APCS		231	8.76%
	Pedestrian crossing zone 2		Zone in which pedestrian finished crossing		
		Within stripes		2275	86.24%
		Between stripes and yield line		165	6.25%
		Between yield line and APCS		198	7.51%
	Pedestrians activities before crossing		Pedestrians activities before crossing		
		Normal		2052	77.79%
		Holding/carrying stuffs		227	8.61%
		Pushing stuffs (bag, stroller, cart)		107	4.06%
		Riding bike		202	7.66%
		On phone		50	1.90%
	Pedestrians activities when crossing		Pedestrians activities when crossing		
		Normal		1726	65.43%
		Holding/carrying stuffs		221	8.38%
		Pushing stuffs (bag, stroller, cart)		138	5.23%
		Riding bike		522	19.79%
		On phone		31	1.18%
	Gender		Gender of pedestrians in a single crossing incident		
		Mixed genders		443	16.79%
		Females only		641	24.30%
		Males only		1554	58.91%
	Pedestrians' age		Approximated age of pedestrians in a single crossing incident		
		Mixed ages		253	9.59%
		Children and teens only		306	11.60%
		Young adults only		512	19.41%
		Adults only		1462	55.42%
		Elderly only		105	3.98%
	Number of pedestrians		Number of pedestrians per crossing incidence		
		One		1873	71.00%
		Two		571	21.65%
		Three or more		194	7.35%
Binary	First to arrive at crosswalk		Whether pedestrian was first to arrive	2070	78.47%
	Pedestrian to/from the bus		Whether a pedestrian was coming/going to the bus	305	11.57%
	Approach from far side		The side pedestrian approached the crosswalk	240	9.10%

4.4.5. Traffic-related variables

The traffic-related variables describe the traffic conditions when pedestrians arrived at the crosswalks. The conditions include: the speeds and number of incoming vehicles; the positions of the front vehicles when pedestrians arrived at the crosswalk; and the AADT (Table 15).

Table 15. Traffic related variables

	Variable type and name	Description	Count	Percent
Categorical	AADT (vpd)			
	Less than 30,000	Annual Average Daily Traffic	1032	39.12%
	Between 30,000 - 40,000		755	28.62%
	Above 40,000		851	32.26%
	Vehicle's position			
	At the stripes	Front vehicle's position from marked stripes	2019	38.27%
	Within 40 ft		725	13.74%
	Between 40ft and 80ft		559	10.60%
	Beyond 80 ft		1973	37.40%
	Vehicles within ECD			
	Few (less than five)	Number of vehicles within effective crosswalk distance	915	17.34%
	Medium (five to ten)		2550	48.33%
	Platoon (ten or more)		1811	34.33%
	Incoming vehicle speed			
	No/stopped vehicles	Speed of the incoming vehicles	1299	24.62%
	Less than 35		644	12.21%
	Between 35 and 45		1317	24.96%
	Greater than 45		2016	38.21%

For the case of AADT, there was no significant difference across the observation compositions, as most of the observations (39.12%) were recorded at crosswalks with AADTs below 30,000 vehicles per day, while the least amount of observations (28.62%) were at locations with AADT between 30,000 and 40,000 vehicles per day. The statistics for vehicles' positions at the moment the pedestrians arrived at the crosswalks show that in most cases, pedestrians arrived while vehicles were either at the stripes (38.27%) or beyond 80ft (37.40%). Mostly, vehicles were travelling at speeds greater than 45mph (38.21%), while the stopping/no vehicles present at the time pedestrians wanted to cross constitutes only about 25%. The predominant number of vehicle within the effective crosswalk distance (ECD) was medium (five to ten vehicles), with about 48% of all observations.

4.4.6. Temporal related variables

The observation distribution according to the time of data collection revealed that the afternoon session had a relatively large number of observations compared to the morning and evening times (Table 16). These times correspond to different trip characteristics such as home to work, work to lunch, work to home, etc.

Table 16. Temporal related variables

	Variable type and name	Description	Count	Percent
Categorical	Time of the day	Time of the day when observations were collected		
	7:00am-11:00am		639	24.22%
	11:00am-1:00pm		283	10.73%
	1:00pm-4:00pm		811	30.74%
	4:00pm-6:00pm		658	24.94%
	6:00pm-9:00pm		247	9.36%

CHAPTER 5: MODEL RESULTS AND DISCUSSION

This chapter presents the results and discussions for the developed models for various interactions at signalized crosswalks. It first avails the pedestrian-driver interaction models, whereby hazard based models (HBMs) and logistic regressions (LR) are presented. Multistate models, which are a family of hazard-based models, were used in this study. The multistate models, which describe the temporal yielding compliance factors explored, are associated with transitions from non-yield to either partial-yield events or full-yield events, and partial-yield events to full-yield compliance. Moreover, the partial-yield, as well as their relationship with near-crash events, are modeled using rare event logistic regressions (RELRs). In addition, logistic regressions that describe not only whether a driver yielded to pedestrians, but also whether the yielding occurred in the designated zone are presented. Lastly, the chapter presents the models for pedestrian-infrastructure interactions. These interactions include the use of pushbuttons, spatial crossing compliance, and temporal crossing compliance, which are all modeled by logistic regression. Moreover, at the crosswalks with TCSs and PHBs, the spatiotemporal crossing compliance was evaluated, whereby multinomial logit, ordered logit, and logistic regression were used. While consideration for the direction of traffic flow was important for most of the models involving transitions of traffic flow, the models that assessed pedestrian behavior did not consider the direction of traffic flow because the subjects in question were not vehicles, but pedestrians. At least a 90% confidence level was considered; however, in rare cases, other statistically insignificant variables, which were important to particular models, were included. It should be noted that due to the similarities in the configurations of CRFBs and CFBs, as well as the small number of observations for CRFBs, the observations from CFBs and CRFBs were combined during modeling.

5.1.Models for driver-pedestrian interactions

As described earlier, the pedestrian-driver interactions were assessed in terms of: 1) the time taken and the transitional states involved when drivers were yielding the right-of-way to pedestrians; 2) the chances for near-miss and partial-yield events; and 3) whether drivers yielded before crossing the advanced yield lines at the crosswalks. The time taken to yield right of way, as well as the transitional states involved, were modeled using the multistate models. The chances for partial-yield and near-miss event occurrence were modeled using Traditional Logistic Regression (TLR) and Rare Event Logistic Regression (RELR). The yielding distance, with respect to the advanced yield line, was modeled using Traditional Logistic Regression (TLR).

5.1.1. Drivers' yielding compliance multistate models

This study applied multistate models, which are the family of hazard-based models, to estimate the transition intensities from one state to another, given that an initial state has been occupied by certain time. Three states – non-yield, partial-yield, and full yield – were observed. Since traffic in two directions of flow behave differently, meaning that the time taken to yield for one direction is different from the second direction of travel, the modeling considered the direction of travel. Moreover, since the yielding for TCSs and PHBs are essentially mandatory, these signal types were used as the bases for comparison to other flash-based signal types; results are presented in Table 18. Meanwhile, the results for the model that involved flash-based signal types are only presented in Table 17. Several traffic, crosswalk, and pedestrian-related factors were used as dependent variables to explain variations in transitions from the non-yielding to yielding of the drivers, as well as for the intermediate step. Several models were developed, and the final model that best fit the data was selected.

As discussed earlier, the model interpretation is based on the hazard ratio, which is defined as the ratio of risk of outcome in the intervention group over the risk of outcome in the control group at a given interval of time, assuming that the subject in study has survived for a certain time. Basically, the hazard ratio is computed as e^{β}, whereby β stands for an estimate of a variable's coefficient. For categorical explanatory variables, the hazard ratio is the ratio of the estimated hazard of the category of interest to that of the base category. On the other hand, for continuous explanatory variables, the hazard ratio is the ratio of the estimated hazard due to the increase of one unit of that variable. If the intervention variable has a hazard ratio greater than one, it implies that the presence of that intervention variable is associated with an increase of the hazard, thus, the decrease in survival. This means that the event of interest will occur faster in the presence of the intervention variable compared to the control variable. Conversely, a hazard ratio of less than one is associated with a decrease of the hazard in the presence of that variable, which translates into a long duration to event of interest. However, if the hazard ratio is one, then there is no significant effect due to the covariate (Spruance et al. 2004). In this study, the events of interest are the transitions of the traffic flow from non-yield to full-yield, non-yield to partial-yield, and partial-yield to full-yield. A covariate that produces a hazard ratio greater than one is desirable, due to the shortened duration.

Table 17 and Table 18 present the model results for flash-based only signal types and all signal types, respectively. Referring to Table 17, the total number of observations was 3976, of which 2229 observations had direct transitions from non-yield to full-yield states, and 221 observations saw the partial-yield state. On the other hand, the number of observations in which vehicles remained in non-yield states were 1526. When considering all signal types (Table 18), similar details were observed. The total number of observations was 5276, whereby 3344 saw the direct transition to full-yield from non-yield, 231 observations saw partial-yield, and 1701 events remained in non-yield states. The discussion of the impact of the explanatory variables is based on flash-based signal types (Table 17) and is divided into four main categories of variables: crosswalk, traffic, and pedestrian-related variables, as well as temporal factors. The discussion comparing models for flash-based signal types (Table 17) and all signal types (Table 18) is briefly presented at the end of this section.

The results in Table 17 show that, in comparison to crosswalks with CFBs and CRFBs, RRFBs are statistically significantly different, at a 95% confidence level, for almost all transitions, except that the RRFBs at stripes and APCSs were not statistically significant for non-yield to partial-yield transitions. Both regular RRFBs and those that have flashing lights on the APCSs are associated with low hazard ratios for all transitions. The implication for the low hazard ratio for this case is that, traffic at these signal types takes more time to yield as compared to CFBs. On average, the time taken for non-yield to full-yield transitions for CFBs and CRFBs is 0.391 times that of RRFBs at stripes only, and 0.653 times that of regular RRFBs with flashing lights at the APCSs. Similar observations can be deduced for non-yield to partial-yield, as the time taken for these transitions at crosswalks with RRFBs at stripes only is 0.237 times that of crosswalks with CFBs and CRFBs. The hazard ratios are much lower for partial-yield to full-yield transitions. The time taken for this transition is 0.009 and 0.073 times that of crosswalks with RRFBs at stripes only and RRFBs at stripes and APCSs, respectively, as compared to crosswalks with CFBs and CRFBs. This finding is conversely to one of previous studies (Fitzpatrick, Potts, et al. 2015) which found no statistical significant difference between RRFBs and CFBs.

Table 17. Multistate-model results for yielding compliance for flashers

	Non-yield to full yield				Non-yield to partial-yield				Partial-yield to full yield			
	Coef	**HR**	**z-stat**	**P-value**	**Coef**	**HR**	**z-stat**	**P-value**	**Coef**	**HR**	**z-stat**	**P-value**
Crosswalk characteristics												
Signal type												
RRFBs at stripes only	-0.939	0.391	-6.482	0.000	-1.440	0.237	-1.972	0.049	-4.763	0.009	-3.638	0.000
RRFBs at stripes and APCSs	-0.427	0.653	-2.694	0.007	-0.651	0.522	-1.019	0.308	-2.622	0.073	-3.056	0.002
Cross stages												
Optional one/two	0.867	2.381	5.424	0.000	0.782	2.186	1.383	0.167	2.404	11.067	3.027	0.002
Strictly one	0.258	1.294	2.403	0.016	-0.398	0.672	-1.222	0.222	-1.092	0.335	-2.798	0.005
Number of lanes												
Six lanes	-0.738	0.478	-5.540	0.000	-0.271	0.763	-0.543	0.587	-1.048	0.351	-1.839	0.066
Seven lanes	-0.861	0.423	-6.951	0.000	-0.784	0.456	-1.584	0.113	-0.898	0.407	-1.699	0.089
Eight to ten lanes	-1.221	0.295	-8.147	0.000	0.153	1.166	0.274	0.784	1.407	4.083	1.939	0.053
Yield here sign	-0.656	0.519	-7.024	0.000	-0.818	0.441	-2.710	0.007	-1.093	0.335	-3.115	0.002
State law sign	-0.497	0.609	-3.388	0.001	-0.014	0.986	-0.024	0.981	0.270	1.310	0.441	0.659
Traffic characteristics												
Vehicles within ECD												
Medium (five to ten)	0.862	2.368	6.486	0.000	1.552	4.720	2.006	0.045	0.254	1.289	0.295	0.768
Platoon (ten or more)	0.440	1.553	3.224	0.001	2.580	13.191	3.347	0.001	0.491	1.634	0.589	0.556
Incoming vehicle speed (mph)												
Less than 35	-0.472	0.624	-4.216	0.000	-0.684	0.505	-1.916	0.055	0.402	1.496	1.005	0.315
Between 35 and 45	-0.760	0.467	-6.834	0.000	-0.337	0.714	-1.002	0.316	0.056	1.058	0.142	0.887
Greater than 45	-0.930	0.395	-8.859	0.000	-0.619	0.539	-2.009	0.045	-0.066	0.936	-0.176	0.860
AADT												
30,000 - 40,000	0.700	2.013	3.551	0.000	0.200	1.222	0.291	0.771	1.044	2.841	1.408	0.159
Above 40,000	0.379	1.461	1.697	0.090	-0.213	0.808	-0.272	0.786	-0.405	0.667	-0.439	0.661
Vehicle's position												
Within 40ft	-0.130	0.878	-1.449	0.147	-0.369	0.691	-1.278	0.201	-0.696	0.499	-2.091	0.037
Between 40ft and 80ft	0.149	1.161	1.693	0.090	-0.384	0.681	-1.176	0.240	-0.131	0.877	-0.421	0.674
Beyond 80ft	-0.084	0.919	-1.274	0.202	-0.379	0.684	-1.827	0.068	-0.154	0.857	-0.703	0.482
Immediate direction	0.949	2.584	17.667	0.000	1.042	2.834	6.968	0.000	0.616	1.852	3.565	0.000

Table 17 Continues

	Non-yield to full yield				**Non-yield to partial-yield**				**Partial-yield to full yield**			
	Coef	**HR**	**z-stat**	**P-value**	**Coef**	**HR**	**z-stat**	**P-value**	**Coef**	**HR**	**z-stat**	**P-value**
Pedestrians, characteristics												
Number of pedestrians												
Two	-0.104	0.902	-1.694	0.090	-0.087	0.916	-0.470	0.638	-0.814	0.443	-3.496	0.000
Three or more	-0.187	0.830	-2.041	0.041	0.338	1.402	1.322	0.186	-0.320	0.726	-1.196	0.232
Pedestrian crossing zone												
Between stripes and yield line	0.359	1.431	3.127	0.002	0.057	1.058	0.133	0.895	0.536	1.710	1.183	0.237
Between yield line and APCS	-0.118	0.888	-0.749	0.454	-0.242	0.785	-0.465	0.642	-0.885	0.413	-1.619	0.106
Model summary												
	n= 2229				n= 2229				n= 221			
	LR test= 598 , 24 df, p<0.001				LR test= 180 , 24 df, p<0.001				LR test= 90 , 24 df, p<0.001			
	-2*Log-likelihood= 25861				-2*Log-likelihood= 2915				-2*Log-likelihood= 1882			

Number of individuals experiencing the intermediate event: 231

Number of events for the direct transition from state 1 to state 3: 2229

Number of individuals remaining in state 1: 1524

Number of events on transition from state 2: 231

Number of censored observations on transition from state 2: 0

Table 18. Multistate-model results for yielding compliance for all signal types

	Non-yield to full yield				Non-yield to partial-yield				Partial-yield to full yield			
Covariates	**Coef**	**HR**	**z-stat**	**P-value**	**Coef**	**HR**	**z-stat**	**P-value**	**Coef**	**HR**	**z-stat**	**P-value**
Crosswalk characteristics												
Signal type												
CFBs & CRFBs	1.864	6.451	16.663	0.000	5.448	232.322	7.718	0.000	2.392	10.935	2.709	0.007
RRFBs	2.145	8.542	25.789	0.000	4.691	108.923	9.874	0.000	1.244	3.471	2.259	0.024
Cross stages												
Optional one/two	0.428	1.534	5.901	0.000	0.191	1.210	0.526	0.599	0.535	1.708	0.988	0.323
Strictly one	0.337	1.401	4.255	0.000	-0.371	0.690	-1.181	0.238	-0.471	0.624	-1.290	0.197
Number of lanes												
Six lanes	-0.028	0.972	-0.258	0.796	-0.400	0.671	-0.819	0.413	-1.062	0.346	-1.825	0.068
Seven lanes	0.059	1.061	0.592	0.554	-0.570	0.566	-1.216	0.224	-0.318	0.728	-0.610	0.542
Eight to ten lanes	-0.390	0.677	-4.043	0.000	0.041	1.042	0.082	0.935	0.324	1.382	0.515	0.606
Yield here sign	-0.221	0.802	-4.183	0.000	-0.780	0.458	-3.137	0.002	-0.700	0.497	-2.183	0.029
State law sign	0.198	1.219	1.712	0.087	-0.047	0.954	-0.087	0.931	0.517	1.677	0.851	0.395
Traffic characteristics												
Vehicles within ECD												
Medium (five to ten)	0.300	1.350	3.636	0.000	1.733	5.660	2.266	0.023	0.103	1.109	0.120	0.904
Platoon (ten or more)	0.107	1.113	1.227	0.220	2.746	15.579	3.606	0.000	0.278	1.320	0.334	0.738
Incoming vehicle speed (mph)												
Less than 35	0.048	1.049	0.569	0.569	-0.851	0.427	-2.497	0.013	0.586	1.798	1.517	0.129
Between 35 and 45	-0.263	0.769	-3.457	0.001	-0.516	0.597	-1.634	0.102	0.221	1.247	0.577	0.564
Greater than 45	-0.285	0.752	-3.824	0.000	-0.723	0.485	-2.514	0.012	0.189	1.208	0.535	0.593
AADT												
30,000 - 40,000	-0.160	0.852	-1.479	0.139	0.521	1.683	0.882	0.378	0.973	2.645	1.333	0.183
Above 40,000	-0.335	0.716	-3.642	0.000	0.106	1.112	0.169	0.866	-0.156	0.856	-0.174	0.862
Vehicle’s position												
Within 40ft	-0.240	0.787	-3.672	0.000	-0.468	0.626	-1.657	0.098	-0.342	0.710	-1.066	0.286
Between 40ft and 80ft	0.046	1.047	0.646	0.518	-0.397	0.672	-1.248	0.212	-0.106	0.900	-0.336	0.737
Beyond 80ft	0.096	1.101	1.722	0.085	-0.390	0.677	-1.921	0.055	0.007	1.007	0.034	0.973
Immediate direction	0.446	1.562	10.792	0.000	0.957	2.604	6.567	0.000	0.626	1.870	3.705	0.000

Table 18 Continues

	Non-yield to full yield				Non-yield to partial-yield				Partial-yield to full yield			
Covariates	**Coef**	**HR**	**z-stat**	**P-value**	**Coef**	**HR**	**z-stat**	**P-value**	**Coef**	**HR**	**z-stat**	**P-value**
Pedestrians characteristics												
Number of pedestrians												
Two	0.037	1.037	0.780	0.436	-0.101	0.904	-0.558	0.577	-0.789	0.454	-3.534	0.000
Three or more	-0.130	0.878	-1.842	0.066	0.284	1.329	1.155	0.248	-0.296	0.744	-1.143	0.253
Pedestrian crossing zone												
Between stripes and yield line	0.346	1.413	3.534	0.000	0.017	1.017	0.039	0.969	0.577	1.781	1.266	0.206
Between yield line and APCS	-0.443	0.642	-3.647	0.000	-0.315	0.730	-0.608	0.543	-0.736	0.479	-1.364	0.173

Model summary

Non-yield to full yield	Non-yield to partial-yield	Partial-yield to full yield
n= 3344	n= 3344	n= 231
LR test= 2120 , 24 df, p<0.001	LR test= 474 , 24 df, p<0.001	LR test= 86 , 24 df, p<0.001
-2*Log-likelihood= 39662	-2*Log-likelihood= 3065	-2*Log-likelihood= 1973

Number of individuals experiencing the intermediate event: 231

Number of events for the direct transition from state 1 to state 3: 3344

Number of individuals remaining in state 1: 1701

Number of events on transition from state 2: 231

Number of censored observations on transition from state 2: 0

Cross stage is another crosswalk related feature that has shown a statistically significant influence on yielding compliance. Compared to strictly two-stage crossings, strictly one-stage and optional one/two stage crossings are associated with a short amount of time for non-yield to full-yield transitions. In fact, the time taken for vehicles to transition from non-yield to full-yield at strictly two stage crosswalks is about 2.4 times of that at optional one/two cross stages, and about 1.3 times the time taken for strictly one-stage crosswalks. Further, the transition from non-yield to partial-yield takes about 0.672 times more for two stages as compared to a one-stage crosswalk, and the transition from partial-yield to full-yield takes about 2.2 times the time taken for two stages, compared to a one-stage crosswalk. However, both one and two cross stage crosswalks were found to be not statistically significantly different for the non-yield to partial-yield transitions. Additionally, for the case of partial-yield to full yield, mixed results could be observed. Traffic at strictly one stage crosswalks took a long time to undergo this transition (hazard ratio =0.335), while at one/two stages crosswalks traffic took less time, which was approximated to be 11 times faster than that of crosswalks with CFBs and CRFBs.

The available number of travel lanes at the crosswalk was also statistically significantly different, at a 95% confidence level for non-yield to full-yield, and partial-yield to full-yield transitions. The non-yield to partial-yield transition was statistically significant, at a 90% level, for only a high number of lanes. The greater the number of lanes, the longer the time taken for transition from non-yield to full-yield. This is to say, for transitions from non-yield to full-yield, traffic at crosswalks with a five-lane roadway took 0.478, 0.423, and 0.295 times the duration of traffic at crosswalks with six, seven, and eight to ten lane roadways, respectively. For the case of non-yield to partial-yield, this variable was not statistically significantly different from zero, at a 95% confidence level (p-value >0.05). For partial-yield to full-yield transitions, there was no clear pattern. The partial-yield to full-yield transition took less time for eight to ten lanes, as compared to a four-lane roadway (HR=4.083). Meanwhile, traffic flow took more time for partial-yield to full-yield transitions at crosswalks with either six or seven lane roadways. The time taken for this transition was lower by a factor of 0.351 and 0.407, as compared to a five-lane roadway.

Two types of signage, “YIELD HERE TO PEDESTRIANS” and “STATE LAW YIELD TO PEDESTRIANS,” were found to be statistically significant at least for one transition. The “STATE LAW YIELD TO PEDESTRIANS” signs were statistically significantly associated with the non-yield to full-yield transition only, while “YIELD HERE TO PEDESTRIANS” signs were associated with all three transitions. Crosswalks with “YIELD HERE TO PEDESTRIANS” signs had low hazard ratios (0.519) for non-yield to full-yield transitions, which means that vehicles at crosswalks with these signs took more time to stop, as compared to the ones that did not have such signs. A similar pattern was observed for the other two transitions. However, since this signage instructs the driver to yield at that specific location, further analysis is required to assess whether the yielding drivers complied to yield at that specific location. Similarly, the “STATE LAW YIELD TO PEDESTRIANS” signs have low hazard ratios for non-yield to full-yield transitions (0.609) and non-yield to partial-yield transitions (0.986), which means that traffic in locations with such signage took more time to yield to pedestrians, as compared to the locations that had no signage. The possible explanation for such an observation may be that the signs were placed after poor yielding compliance was observed, aiming at providing more information to drivers. However, with these findings, it can be concluded that both “STATE LAW” and “YIELD HERE” signs have no positive influence on yielding compliance.

The incoming vehicle speeds variable was associated with a long time for drivers to yield. This variable was statistically significant, at a 95% level for non-yield to full-yield and non-yield to partial-yield transitions. For both transitions, the variable showed a low hazard ratio, which implies that traffic takes a longer time to yield when the incoming traffic volume is high. For instance, for non-yield to full-yield transitions, the incoming speeds of less than 35 mph, between 35 mph and 45mph, and greater than 45mph had hazard ratios of 0.602, 0.454, and 0.385, respectively.

Moreover, number of vehicles within an effective crosswalk distance (ECD) is associated with a short time for drivers to yield. This variable had positive coefficients and hazard ratios greater than one for all three transitions, but it was statistically significant, at a 95% confidence level for non-yield to full-yield and non-yield to partial-yield transitions only. The more vehicles present in the ECD the shorter time it took for the non-yield to full-yield and non-yield to partial-yield transitions to occur. In fact, the non-yield to partial-yield transitions took about 13 times less time to occur when there were ten or more vehicles in the ECD, as compared to when there were less than five vehicles. A similar trend was observed for non-yield to full-yield transitions; however, the magnitude was high for a fewer number of vehicles.

For two-way roadways, two directions of traffic flow are available: the immediate direction and the farther direction. The immediate direction is the direction of traffic flow that pedestrians first encounter. Table 17 shows that vehicles in the immediate direction took 2.584, 2.834, and 1.852 times less time to change from non-yield to full-yield, non-yield to partial-yield, and partial-yield to full-yield, respectively, as compared to vehicles in the second/farther direction of flow.

The AADT factor was found to have mixed results. However, this variable was statistically significant for non-yield to full-yield transitions only. For this transition type, the variable depicted a positive association. The hazard ratios for non-yield to full-yield transitions were found to be 2.013 for the AADT ranging between 30,000 vpd and 40,000 vpd, and 1.461 for the AADT above 40,000 vpd, which implies that less time is taken for this transition. This observation is counterintuitive, since it is expected that the higher the AADT, the longer it will take for non-yield to full-yield transitions.

The positions of vehicles when pedestrians arrived at crosswalks were also investigated and associated with the transitions to yield. Mixed results were observed for this variable. The situations in which pedestrians arrived at a crosswalk when vehicles were within 40ft from the stripes, were associated with long durations for the transitions that involved partial-yield. Further, the transition directly to a full-yield was observed to take long time if vehicles were either too close or too far away from the crosswalk. To be exact, the hazard ratio for non-yield to full-yield transitions was found to be 1.161 if the vehicles were between 40ft-80ft from the marked stripes, 0.878 if the vehicles were within 40ft from the stripes, and 0.919 if the vehicles were beyond 80ft from the stripes. For non-yield to partial-yield, as well as partial-yield to full-yield, only situations in which pedestrians arrived while the vehicles were beyond 80ft and between 40ft-80ft from the marked stripes were found to be statistically significant, at a 90% significant level, with hazard ratios of 0.684 and 0.499, respectively.

The number of pedestrians and the zones in which the pedestrians crossed were found to associate to the time taken for drivers to yield. The presence of more pedestrians waiting to cross was statistically significantly associated with an increased duration for non-yield to full-yield and partial-yield to full-yield transitions, while the association with non-yield to a partial-yield transition was observed to be statistically insignificant. The crossing zone, on the other hand, was found to be associated with non-yield to full-yield transitions only. If pedestrians crossed between the stripes and yield line, the traffic flow transition from non-yield to full-yield resulted in a hazard ratio of 1.431, while the same crossing location resulted into hazard ratios of 1.058 and 1.710 for non-yield to partial-yield and partial-yield to full-yield, respectively. On the other hand, if pedestrians crossed between the yield line and APCSs, the hazard ratio for partial-yield to full-yield became 0.888 for non-yield to full-yield transitions. This implies that crossing away from the marked stripes is associated with a long duration for transitions, although they are statistically insignificant.

A comparison between the models developed on data collected from flash-based only signal types and from all signal types revealed significant differences. The first major difference was that compared to TCSs and PHBs signals, flash-based signals had higher hazard ratios for all three transitions, which means they had short transition durations. This is of no surprise since it is according to their design. The second major difference was the complete change of the coefficients' signs for the number of lanes and AADT variables. Moreover, some other categories of the variables were observed to have changes in their coefficients' signs and significance levels (p-values). Lastly, a decrease and increase of the magnitudes of the coefficients for different variables was observed.

5.1.1.1.Comparison of the yielding compliance models

The previous studies in this area have implemented binary and count models for evaluating the yielding compliance of drivers. This study presents a binary model (Table 19) and a vehicle count model (Table 20) to make a comparison with the developed multistate models.

Table 19. Logistic regression results for drivers' yielding compliance

	Flashers only				**All signals**			
Drivers yielded=1	**Coef**	**OR**	**z-stat**	**P-value**	**Coef**	**OR**	**z-stat**	**P-value**
Crosswalk characteristics								
Signal type								
RRFBs at stripes only	0.770	2.160	3.530	0.000				
RRFBs at stripes and APCSs	0.736	2.088	2.787	0.005				
CFBs & CRFBs					-1.782	0.168	-8.053	0.000
RRFBs					-0.522	0.594	-3.349	0.001
Cross stages								
One/two	1.993	7.339	6.638	0.000	2.220	9.208	14.296	0.000
One	0.598	1.818	3.316	0.001	1.299	3.667	8.878	0.000
Number of lanes								
Six lanes	1.574	4.826	6.487	0.000	2.236	9.359	11.135	0.000
Seven lanes	0.462	1.587	2.387	0.017	0.774	2.168	4.831	0.000
Eight to ten lanes	1.248	3.483	4.896	0.000	1.437	4.209	7.885	0.000
Yield here sign	-0.889	0.411	-5.208	0.000	-0.553	0.575	-4.911	0.000
State law sign	0.346	1.414	1.636	0.102	0.374	1.454	2.063	0.039
Traffic characteristics								
Vehicles within ECD								
Few (five to ten)	2.574	13.116	15.027	0.000	-1.782	0.168	-8.053	0.000
Platoon (more than ten)	3.661	38.885	18.889	0.000	-0.522	0.594	-3.349	0.001
Incoming vehicle speed (mph)								
Less than 35	-0.101	0.904	-0.566	0.572	-0.061	0.941	-0.399	0.690
Between 35 and 45	-0.498	0.608	-3.088	0.002	-0.537	0.585	-4.097	0.000
Greater than 45	-0.491	0.612	-3.164	0.002	-0.347	0.707	-2.703	0.007
AADT								
30,000 - 40,000	-0.882	0.414	-2.576	0.010	-1.184	0.306	-5.586	0.000
Above 40,000	-1.934	0.145	-4.879	0.000	-2.261	0.104	-11.00	0.000
Vehicle's position								
Within 40ft	-0.511	0.600	-3.086	0.002	0.043	1.044	0.312	0.755
Between 40ft and 80ft	-0.292	0.747	-1.763	0.078	-0.047	0.954	-0.324	0.746
Beyond 80ft	-0.273	0.761	-2.334	0.020	-0.194	0.824	-1.917	0.055
Immediate direction	-0.802	0.448	-9.105	0.000	-0.647	0.523	-8.433	0.000
Pedestrians characteristics								
Number of pedestrians								
Two	0.352	1.422	3.221	0.001	0.499	1.648	5.280	0.000
Three or more	0.779	2.180	4.346	0.000	0.701	2.016	4.623	0.000
Pedestrian crossing zone								
Between stripes and yield line	-0.346	0.708	-2.047	0.041	-0.185	0.831	-1.212	0.226
Between yield line and APCS	-2.617	0.073	-15.65	0.000	-2.337	0.097	-16.77	0.000
Constant term	-1.379	0.252	-3.792	0.000	-0.659	0.517	-2.311	0.021
Model summary								
Number of observations	3976				5276			
AIC score	3792				4793			
BIC score	3949				4957			

Table 20. Negative Binomial (NB) results for drivers' yielding compliance

	Flashers only				All signals			
Number of vehicles	**Coef**	**IRR**	**z-stat**	**P-value**	**Coef**	**IRR**	**z-stat**	**P-value**
Crosswalk characteristics								
Signal type								
RRFBs at stripes only	1.018	2.767	5.629	0.000				
RRFBs at stripes and APCSs	1.128	3.089	5.838	0.000				
CFBs & CRFBs					3.990	54.062	18.172	0.000
RRFBs					5.033	153.330	25.934	0.000
Cross stages								
One/two	0.646	1.907	3.385	0.001	0.548	1.730	3.761	0.000
One	-0.18	0.835	-1.547	0.122	-0.175	0.839	-1.506	0.132
Number of lanes								
Six lanes	1.492	4.446	8.803	0.000	1.443	4.233	9.562	0.000
Seven lanes	0.791	2.207	5.133	0.000	0.830	2.292	5.808	0.000
Eight to ten lanes	1.685	5.391	9.214	0.000	1.683	5.381	11.249	0.000
Yield here sign	-0.22	0.797	-2.119	0.034	-0.211	0.810	-2.135	0.033
State law sign	0.892	2.440	5.040	0.000	0.868	2.381	5.556	0.000
Traffic characteristics								
Vehicles within ECD								
Few (five to ten)	0.455	1.577	3.255	0.001	0.379	1.460	2.786	0.005
Platoon (more than ten)	1.720	5.582	11.989	0.000	1.601	4.956	11.471	0.000
Incoming vehicle speed (mph)								
Less than 35	0.934	2.544	7.413	0.000	0.950	2.585	7.611	0.000
Between 35 and 45	1.313	3.719	11.122	0.000	1.317	3.733	11.295	0.000
Greater than 45	1.653	5.223	14.645	0.000	1.649	5.201	14.782	0.000
AADT								
30,000 - 40,000	-1.54	0.215	-6.552	0.000	-1.435	0.238	-8.466	0.000
Above 40,000	-2.03	0.132	-7.584	0.000	-1.891	0.151	-10.23	0.000
Vehicle's position								
Within 40ft	-0.33	0.721	-3.300	0.001	-0.359	0.699	-3.677	0.000
Between 40ft and 80ft	-0.72	0.485	-7.060	0.000	-0.739	0.478	-7.319	0.000
Beyond 80ft	-0.63	0.535	-8.245	0.000	-0.629	0.533	-8.312	0.000
Immediate direction	-0.69	0.502	-12.14	0.000	-0.713	0.490	-12.64	0.000
Pedestrians characteristics								
Number of pedestrians								
Two	-0.03	0.968	-0.461	0.645	-0.026	0.974	-0.373	0.709
Three or more	0.183	1.200	1.608	0.108	0.159	1.172	1.408	0.159
Pedestrian crossing zone								
Between stripes and yield line	-0.28	0.756	-2.282	0.022	-0.217	0.805	-1.792	0.073
Between yield line and APCS	-0.91	0.403	-7.805	0.000	-0.842	0.431	-7.334	0.000
Constant term	-1.65	0.192	-6.263	0.000	-5.577	0.004	-22.49	0.000
				Model summary				
Number of observations								
AIC score		12083				12518		
BIC score		12246				12688		
Log-likelihood		-6015				-6233		

It can be observed that, in comparison to the multistate models developed in this study, the binary and count models can be only comparable to the non-yield to full-yield transitions of the multistate model, since none of the binary and vehicle count models included a partial-yield state. In addition, apart from ignoring partial-yield, the binary models also ignore the time taken for drivers to stop for pedestrians. On the other hand, the vehicle counts models consider the number of non-yielding vehicles until drivers stop for pedestrians. Moreover, the signs of the coefficients for all three models vary significantly.

5.1.2. Partial-yield incidents and near-miss events models

In this section, the TLR and RELR model results, as well as discussions for partial-yield and near-miss events, are presented. A 95% confidence level was adopted for this study; therefore, for a variable to be included in the model, at least one of its categories should have a p-value of 0.05 or less. Due to this condition, there are variables that are observed in the partial-yield event models, but not in the near-miss event models. Moreover, the "YIELD HERE" and "STATE LAW" variables were not found to be statistically significant, at a 95% confidence level. This discussion first presents the comparison between TLR and RELR, followed by the discussion for partial-yield incidents, with the discussion for near-miss events presented last.

The comparison between TLR and RELR reveals that the coefficients for TLR are slightly larger in magnitude than those for RELR. The difference in coefficient magnitude is larger for the near-miss events models compared to partial-yield incidents models. This can be explained by the fact that the percentage composition for near-miss events (1.6%) is less than that of partial-yield events (5.6%), which makes near-miss events rarer than partial-yield events. Although there were observed differences in coefficient magnitudes, the general goodness of fit parameters for both partial-yield incidents and near-miss events models reveal no difference. The resulting AIC for partial-yield incidents models was found to be 1433 for both TLR and RELR, while that of near-miss events was 504. Based on these AIC scores, both TLR and RELR had similar performances in fitting the data. However, the results' discussions are based on RELR, since the TLR tends to underestimate the probability of event occurrence, as reported by King and Zeng (King and Zeng 2001).

5.1.2.1.Model results for partial-yield incidents

Presented in Table 21 are the best-fit TLR and RELR models for partial-yield incidents. The discussion is based on the odds ratios, and is subdivided into crosswalk-related, traffic-related, and pedestrian-related variables.

Crosswalk-related attributes, which include signal type, number of lanes, and median type, as well as signal status at the next intersection downstream of the crosswalk, have shown both positive and negative associations to partial-yield events. According to the results in Table 21, it can be observed that compared to CFBs and CRFBs, RRFBs were negatively associated to partial-yield events. The odds of partial-yield events at RRFBs that are positioned at marked stripes were as low as 0.446. This is contrary to a study (Fitzpatrick, Potts, et al. 2015) that found no significant difference in driver behaviors due to different shapes of flashing beacons, although the authors' focus was on yielding compliance. The current study also found that CRFBs, and the addition of RRFBs at APCSs, had no statistically significant difference, at a 95% confidence level, compared to CFBs.

Table 21. Model results for partial-yield events

	Rare Event Logistic Regression (RELR)			Traditional Logistic Regression (TLR)		
Pr (partial-yield =1)	**Coef**	**OR**	**P-value**	**Coef**	**OR**	**P-value**
Crosswalk characteristics						
Signal type						
CRFBs	-0.477	0.621	0.247	-0.478	0.6197	0.245
RRFB at stripes only	-0.808	0.446	0.036	-0.813	0.4435	0.034
RRFB at stripes and APCS	-0.560	0.571	0.075	-0.563	0.5695	0.074
Number of lanes						
Six lanes	0.546	1.726	0.140	0.547	1.7278	0.139
Seven lanes	0.031	1.031	0.928	0.029	1.0292	0.932
Eight to ten lanes	0.857	2.357	0.034	0.858	2.3588	0.034
Median type						
Narrow-raised	0.797	2.219	0.004	0.803	2.2311	0.003
Wide-raised	0.533	1.704	0.220	0.540	1.7167	0.214
Green traffic signal	0.526	1.692	0.002	0.530	1.6987	0.002
Traffic characteristics						
Vehicles within ECD						
Few (five to ten)	1.568	4.798	0.040	1.581	4.8597	0.038
Platoon (more than ten)	2.924	18.61	0.000	2.943	18.965	0.000
Incoming vehicle speed (mph)						
Less than 35	0.353	1.423	0.340	0.352	1.4221	0.341
Between 35 and 45	0.705	2.024	0.038	0.706	2.0264	0.038
Greater than 45	0.644	1.905	0.045	0.644	1.9042	0.045
Vehicle's position						
Within 40ft	-0.441	0.643	0.146	-0.440	0.6439	0.147
Between 40ft and 80ft	-0.565	0.568	0.084	-0.569	0.5664	0.082
Beyond 80ft	-0.472	0.623	0.026	-0.475	0.6221	0.026
Vehicles in immediate direction	0.425	1.529	0.008	0.430	1.5375	0.007
Time to yield	0.022	1.022	0.000	0.022	1.0223	0.000
Pedestrian characteristics						
Gender						
Females only	0.604	1.829	0.067	0.608	1.8364	0.065
Males only	0.694	2.001	0.024	0.699	2.0121	0.023
Number of pedestrians						
Two	0.216	1.241	0.347	0.219	1.2451	0.339
Three or more	1.017	2.765	0.003	1.027	2.7939	0.002
Constant	-10.222	4E-05	0.000	-7.405	0.0006	0.000
Model fit parameters						
	Number of observations = 3976					
	AIC = 1433			AIC = 1433		

The remaining crosswalk related factors have shown positive associations to partial-yield incidents. The first is the median type: the odds of a partial-yield occurrence for crosswalks that have a narrow-raised median are 2.219, as compared to locations with no median. The wide-raised type of median has shown no statistically significant difference for partial-yield occurrences, when compared to no median locations. The number of lanes available at a crosswalk is associated with an increased chance of partial-yield events. The greater the number of lanes, the higher the chance of partial-yield occurrence. In fact, crosswalks located on roadways with eight or more lanes have more than two times (2.357) the chance of partial-yield events being observed. The crosswalks on roads where lanes are between four and eight have no statistically significant difference from those located at five-lane roadways. A similar conclusion related to the influence of number of lanes to partial-yield incidents was presented by (M. Mitman, Cooper, and DuBose 2010; M. F. Mitman, Ragland, and Zegeer 2008). For crosswalks located within 0.1 mile upstream of signalized crosswalks, the chance that a partial-yield incident will occur if the traffic signal at the next intersection is green is high, by about 52.6%, as compared to when the signal lights are red.

Apart from crosswalk related factors, traffic factors, which include the number of vehicles within the effective crosswalk distance (ECD), incoming vehicle speed, vehicle position when a pedestrian arrived at a crosswalk, and direction of flow were also evaluated. Results in Table 21 show that the greater the number of vehicles within the ECD, the higher the odds were for partial-yield events. This variable shows that if there were ten or more vehicles within the ECD, the odds were about 19 times more that partial-yield events would occur, as compared to when there were fewer than five vehicles. Comparatively, the odds were about five times that a partial-yield incident would occur if there were five to ten vehicles in the ECD.

In connection with the number of vehicles within the ECD, the zone in which the front vehicle was positioned when a pedestrian arrived at a crosswalk was also evaluated. The results showed that the farther away the front vehicle was when a pedestrian arrived at a crosswalk, the lower the chance of partial-yield incidents. The chance that a partial-yield incident would occur decreased by about 43% and 38% when the front vehicle was located between 40ft and 80ft and beyond 80 ft, respectively. This finding could be included in the Manual on Uniform Traffic Control Devices (MUTCD) (Manual on Uniform Traffic Control Devices (MUTCD) 2009) to support the requirements for positioning yield lines. According to the MUTCD (Manual on Uniform Traffic Control Devices (MUTCD) 2009), the yield line should be neither less than 40ft nor 180ft from the crosswalk signal; however, there are no specific reasons given for these requirements.

Incoming vehicle speeds were also found to have positive associations with partial-yield events. The higher the speeds, the higher the likelihood of partial-yield incident occurrences. That is, if the incoming vehicle speed exceeded 35 mph, the odds of partial-yield events were about double, compared to when the speed was less than that margin. The vehicles in the immediate direction were about 53% more likely to be part of partial-yield events, as compared to those in the farther direction of traffic flow.

Last for traffic-related variables, is the time taken by drivers to yield right-of-way to pedestrians. The crosswalks where time to yield was significantly longer, were more likely to have partial-yield events. For this variable, for every second that vehicles continued to flow while a pedestrian was waiting to cross, the chances of partial-yield events increased by 2.2%.

The number of pedestrians and their gender were the two pedestrian-related factors associated with partial-yield events. The odds for partial-yield events were high for both male only (2.001) and female only (1.829) pedestrians, as compared to when they were mixed (male and female crossing together). Moreover, in the presence of three or more pedestrians in need of crossing, the chances of a partial-yield incident increased by about three times, as compared to when there was only one person waiting to cross. The high likelihood of partial-yield incidents in the presence of a group of people wanting to cross at one time can be explained by the heterogeneity in the pedestrians' actions within a group. In the same group, some pedestrians might want to cross, while others still want to remain on the side, or cross at a different location within a crosswalk.

5.1.2.2.Model results for near-miss events

Table 22 presents the RELR and TLR results for the near-miss events. Several variables were found to associate with near-miss event occurrence. These variables can be categorized as crosswalk-related, traffic-related, temporal factors, and pedestrian-related.

Considering the traffic related factors, the higher the incoming vehicle speed, the higher the odds for near-miss events. For any speed greater than 45mph, the odds were 2.5 higher that a near-miss event might be observed, compared to when speed was less than 45mph. The yielding behavior of drivers also played a vital role in near-miss events. This could be revealed by the association between near-miss and partial-yield events, as well as the time taken to yield. The occurrence of partial-yield events at any crosswalk increased the chance of near-miss occurrences by about 60.2%. Similarly, for each second that elapsed when a pedestrian was waiting to use a crosswalk, the chances of near-miss events increased by about 3.1%.

Table 22. Model results for near-miss events

	Rare Event Logistic Regression (RELR)			Traditional Logistic Regression (TLR)		
Pr (Near-miss=1)	**Coef**	**OR**	**P-value**	**Coef**	**OR**	**P-value**
Crosswalk characteristics						
Cross stages						
Optional one/two	2.264	9.625	0.051	2.345	10.438	0.043
Strictly one	4.343	76.949	0.000	4.473	87.588	0.000
Number of lanes						
Six lanes	3.616	37.202	0.000	3.721	41.287	0.000
Seven lanes	3.649	38.425	0.000	3.747	42.390	0.000
Eight to ten lanes	3.882	48.513	0.000	4.000	54.623	0.000
Land use						
Residential	2.741	15.501	0.000	2.792	16.319	0.000
Commercial	0.086	1.090	0.910	0.103	1.109	0.892
Yield line to APCS distance						
Between 100 and 200ft	-2.522	0.080	0.000	-2.592	0.075	0.000
More than 200ft	-3.639	0.026	0.000	-3.728	0.024	0.000
In and out turn lanes	1.661	5.266	0.003	1.693	5.437	0.002
Traffic characteristics						
Incoming vehicle speed (mph)						
Less than 35	-0.502	0.606	0.470	-0.518	0.595	0.455
Between 35 and 45	0.143	1.154	0.793	0.153	1.166	0.778
Greater than 45	0.915	2.497	0.057	0.938	2.555	0.051
Partial-yield event	1.602	4.964	0.000	1.630	5.104	0.000
Time to yield	0.030	1.031	0.000	0.031	1.031	0.000
Pedestrian characteristics						
Number of pedestrians						
Two	0.365	1.441	0.326	0.365	1.440	0.327
Three or more	2.377	10.778	0.000	2.422	11.264	0.000
Temporal factors						
Time of the day						
Between 11:00am and 1:00pm	-0.278	0.757	0.775	-0.294	0.745	0.763
Between 1:00pm and 4:00pm	0.733	2.081	0.240	0.753	2.122	0.227
Between 4:00pm and 6:00pm	2.567	13.030	0.000	2.626	13.814	0.000
Past 6:00pm	2.960	19.291	0.000	3.024	20.574	0.000
Constant	-18.456	0.000	0.000	-17.24	0.000	0.000
		Model fit parameters				
			Number of observations= 3976			
		AIC = 504			AIC = 504	

The crosswalk related factors, which include cross stages, number of lanes, turning lanes, and land use, were found to have a very high association with near-miss events. It was revealed that the chances for near-miss events for strictly one-stage, and optional one/two stage crosswalks were about 77 and 10 times higher respectively, compared to those of strictly two-stage crosswalks. Apart from that, the greater the number of lanes at the crosswalk, the higher the possibility of near-miss events. The odds for near-miss events for six, seven, and eight to ten lane roadways were 37.2, 38.4, and 48.5 higher, respectively, as compared to a five-lane roadway. The possible reason for this is the fact that with many lanes, there is a higher chance of having drivers with different aggressive behaviors, which tends to impact the way traffic flow comes to a complete stop. The presence of inside and outside turning lanes at a crosswalk also showed an association with near-miss events. Table 22 shows that the odds for near-miss events were about five times greater when there were turning lanes at crosswalks. Moreover, the distance between yield lines and APCSs played a great role in near-miss event occurrences. The wider the distance, the lower the chances of near-misses, as revealed by the odds of 0.08 and 0.026 for the distance between 100 and 200 feet and more than 200 ft, respectively. Lastly, residential land use has also shown a significant association with near-miss events. The odds for this variable for near-miss occurrence are about 15 times higher compared to mixed land use.

The pedestrian characteristic that has a significant association with the near-miss events is the number of pedestrians crossing. The higher the number of pedestrians, the higher the odds for near-miss events, as shown by the 10.8 odds for three or more pedestrians crossing, when compared to one pedestrian crossing. The situation can be explained by the heterogeneity in pedestrians' behavior, which may affect drivers' decisions.

Time of the day was also found to be an important factor leading to near-miss event occurrence. Compared to morning hours (6:00am to 11:00am), mixed results on chances of near-miss events for afternoon and evening hours were observed. Table 22 shows that the window between 11:00am and 1:00pm had a low association with near-miss events, while any time after 4:00pm had a high association with near-miss occurrences. The odds for near-miss events past 6:00pm were the highest (19.2) of the entire day. The reason for this observation might be the dark conditions during that time.

5.1.3. Drivers' spatial yielding compliance models

Safe pedestrian-driver interaction requires not only drivers to yield the right-of-way to pedestrians, but also that the yielding should occur in designated zones. Thus, spatial yielding compliance is equally as important as yielding compliance itself. The closer the drivers yield right-of-way to pedestrians, the more danger they pose, and the higher the discomfort to pedestrians. Most crosswalks have a yield line with a "YIELD HERE FOR PEDESTRIANS" sign instructing drivers to yield to pedestrians. Spatial yielding compliance is assessed by determining whether drivers yield before the yield line. Model results presented in Table 23 describe drivers' yielding compliance according to the zones.

5.1.3.1.Models performance results

In general, the model with both crosswalk-related and non-crosswalk-related features performed better than the other two. Considering prediction accuracy, comparing the three models, the model with non-crosswalk-related features was the inferior (79%), followed by the model with crosswalk-related features (81%), while the one with both crosswalk-related and non-crosswalk-related features performed slightly better (82%). The prediction accuracy results suggest that, a crosswalks' features can only predict up to 81% of drivers' spatial yielding compliance, which is only 1% less than the model that has the combined features.

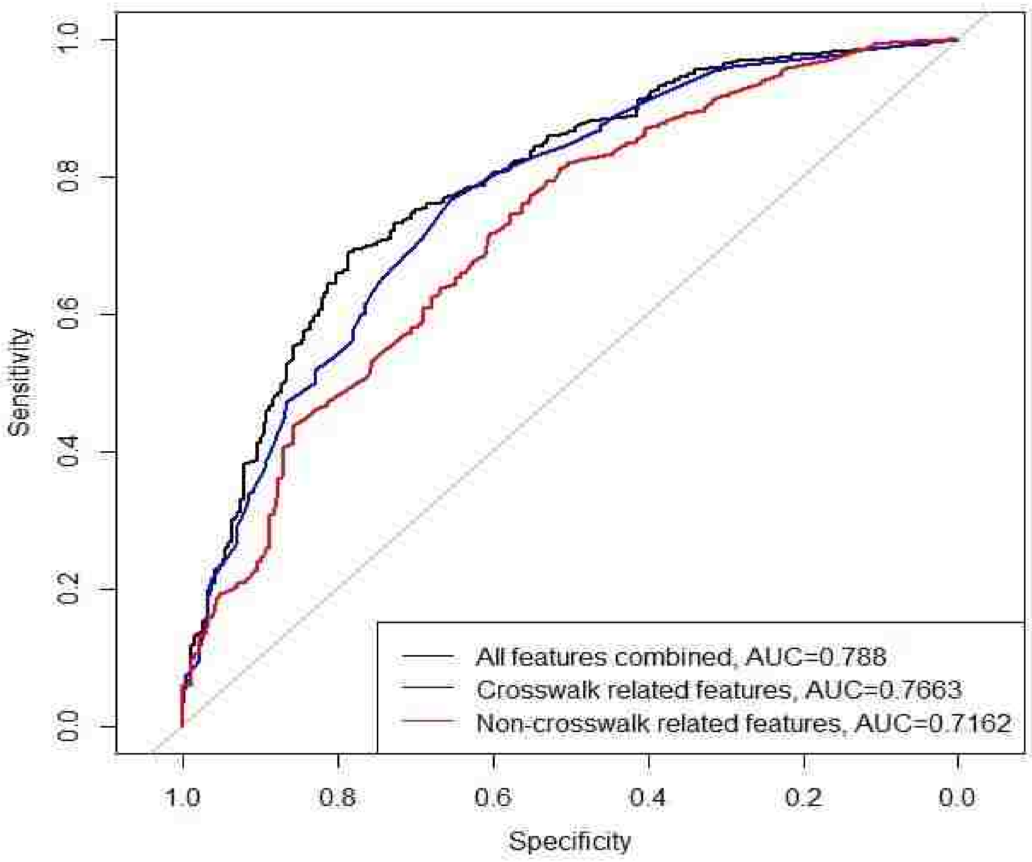


Figure 26. The area under the ROC curve (AUC)

Similarly, the differences between the AUCs for the three competing models was very small (Figure 26). Although the AUC for the model with all features (0.788) was the largest, it differed from the one with crosswalk-related features only by 0.0217. Neither prediction accuracy nor AUC clearly provided the best performing model, as the differences were too small. The BIC and AIC, however, clearly showed that the model that had all features combined performed better than the one with crosswalk-related features alone. This model had a BIC score of 3568, compared to a BIC score of 3585 for the model with crosswalk-related features only. The difference between BIC scores for the models was 17, which is significant (Fabozzi et al. 2014). A similar trend was observed for AIC scores (Table 23).

5.1.3.2.Models results discussion

The models' interpretations are based on the odds ratios (OR), since coefficients do not convey a straightforward meaning for logistic regressions. The variables with an OR of greater than one are positively associated with the occurrence of spatial yielding. On the other hand, an OR of less than one implies that the variable is negatively associated with spatial yielding compliance. Apart from that, for the general performance of the models, the AICs and BICs are used. Based on these scores, it can be observed that the model that combined both crosswalk and non-crosswalk-related features fit the data more perfectly (Table 23). Moreover, for crosswalk-related variables, there was a slight difference in the coefficients/odds ratios for the model with crosswalk-related variables only, to those of combined variables. Thus, the discussion of the results is based on the all features model. However, the comparison between the model with crosswalk-related variables only, and the all features' model is presented. The model results presented in Table 23 are grouped in terms of crosswalk characteristics, traffic characteristics, pedestrian-related factors, and temporal factors.

From Table 23 it can be observed that several crosswalk related factors are associated with the spatial yielding compliance. Considering signal type, the odds for a driver yielding beyond the yield line increases when RRFBs are used. More importantly, the presence of an RRFB at the crosswalk stripes and APCS increases odds of spatial yielding compliance to about 13 times, as compared to an RRFB at the stripes only.

Table 23. Logistic Regression (LR) results for drivers' spatial yield compliance

	Crosswalk features only			All features combined		
Pr(Comply to spatial yield =1)	**Odds Ratio**	**z-statistic**	**P-value**	**Odds Ratio**	**z-statistic**	**P-value**
Crosswalk characteristics						
Signal type						
RRFB at stripes only	1.797	2.057	0.040	1.668	1.633	0.103
RRFB at stripes and APCS	10.043	7.787	0.000	13.246	8.290	0.000
Land use						
Residential	4.816	7.297	0.000	4.543	6.189	0.000
Commercial	27.149	17.426	0.000	23.322	15.124	0.000
Cross stages						
One/two	0.439	-2.383	0.017	0.600	-1.043	0.297
One	0.655	-1.175	0.240	0.203	-3.697	0.000
Yield line to marked stripes distance						
Between 40 and 80 ft	0.543	-4.583	0.000	0.528	-4.433	0.000
More than 80 ft	0.471	-5.388	0.000	0.641	-2.950	0.003
Yield line to APCS distance						
Between 100 and 200 ft	2.176	4.832	0.000	2.2 57	4.420	0.000
More than 200 ft	2.170	3.672	0.000	3.022	4.452	0.000
State's law sign	9.100	12.196	0.000	16.421	11.198	0.000
PED XING sign	0.959	-0.117	0.906	5.458	2.502	0.012
Arrived during active flashes	0.595	-4.676	0.000	0.763	-2.211	0.027
Traffic characteristics						
AADT (vpd)						
Between 30,000 - 40,000				0.279	-5.059	0.000
Above 40,000				0.153	-3.592	0.000
Vehicles in immediate direction				0.428	-12.231	0.000
Pedestrian characteristics						
Pedestrians activities when crossing						
Holding/carrying stuffs				1.353	1.713	0.087
Pushing stuffs (bag, stroller, cart)				0.556	-3.541	0.000
Riding bike				0.824	-1.662	0.097
On phone				5.890	3.123	0.002
Number of pedestrians						
Two				1.079	0.628	0.530
Three or more				0.724	-1.812	0.070
Pedestrian crossing zone						
Between stripes and yield line				1.078	0.443	0.658
Between yield line and APCS				3.612	5.130	0.000
Temporal factors						
Time of the day						
Between 11:00 and 13:00				0.575	-3.596	0.000
Between 13:00 and 16:00				0.880	-0.836	0.403
Between 16:00 and 18:00				0.510	-4.862	0.000
Past 18:00				0.849	-0.807	0.419
Constant	0.756	-0.728	0.467	2.400	1.915	0.056
	Model fit parameters					
	AIC = 3497			AIC = 3386		
	BIC = 3585			BIC = 3568		

This finding shows the safety benefits gained by alerting drivers in advance. The difference between crosswalk-related and all features combined models was only about three-fold.

The land use areas where crosswalks are located affects the spatial yielding compliance. This is exposed by the odds of 4.5 and 23.3 times for residential and commercial land use, respectively, as compared to mixed land use for the model with all features. The model for crosswalk features only had relatively high odds for this variable. The crosswalk configuration, as the contributory factor for spatial yielding compliance, has shown that a single-stage crosswalk had low odds (0.203) of spatial yielding compliance compared to a two-stage crosswalk. The one/two crossing stages had no statistically significant difference from two-stage crossings, at a 95% confidence level.

Another crosswalk related factor that is statistically significantly associated with spatial yielding compliance is the distance between stripes and APCSs. This distance can be grouped into two categories: first is the distance between the stripes and yield line; and second is the distance between the yield line and APCSs. Table 23 shows that as the distances between stripes and yield lines increases, the spatial yielding compliance decreases, while the increase of the distance between yield lines and APCSs results into an increase in spatial yielding compliance. Compared to the yield line to stripes distance that is less than 40 ft, the distance between 40 and 80 ft resulted into a decline in yielding compliance by about 47%, while an increased distance to beyond 80 ft resulted into a decline of the same by 36%. Based on the yield line to APCSs distances, when the distance was between 100 ft and 200 ft, and above 200 ft, the spatial yielding compliance increased by about 2.3 and 3.0 times, respectively, compared to when the distance was less that 100ft. For this factor, there was small change in odds ratio when the model with crosswalk-related variables only was considered (Table 23). The increased spatial yielding compliance as a result of the increased distance from yield line to APCSs can be explained by the longer distance and time presented to drivers to make yielding decisions. Contrarily, an increased distance between the striped and yield line provides less room for drivers to yield before the yield line, as some of them might already be in the zone when pedestrians arrive at the crosswalk.

Further, the signage at crosswalks has demonstrated a great effect in spatial yielding compliance. In fact, the presence of a “STATE LAW” sign resulted into the odds of 16.4 of drivers’ spatial yielding compliance, while the “PED XING” sign has shown 5.4 odds of spatial yielding compliance. If drivers arrived at the crosswalk while the flashes or walk signals were active, the odds were 0.763 that they would adhere to spatial yielding compliance. This is because, drivers might have already passed the yield line, especially for the crosswalks with long spans between the stripes and yield lines.

Traffic characteristics that influence the spatial yielding compliance include the AADT and direction of traffic flow. The higher the AADT, the lower the spatial yielding compliance, as shown by the 0.279 and 0.153 odds on the AADT between 30,000 and 40,000, and above 40,000 respectively. On the other hand, vehicles in the immediate direction of traffic flow were less likely (odds=0.428) to comply with the spatial yielding as compared to the farther direction of traffic flow.

Pedestrian-related factors include the number of pedestrians, crossing zone, activities during crossing, and flashing status when pedestrians arrive at a crosswalk. The number of pedestrians that are using the crosswalk has shown mixed findings. The higher number of pedestrian crossing was associated with less spatial yielding compliance of drivers. That is, when there were three or more, the spatial yielding compliance declined by about 37.6%, as compared to when there was one pedestrian crossing. Contrary, if only two pedestrians were crossing, the odds that drivers would yield within the designated zone increased by 7.9%, although the results are not statistically significant at a 95% level. The zone in which pedestrians cross also played important role in spatial yielding compliance. As pedestrian crossed away from the marked stripes, the odds of spatial yielding increased. Being specific, the odds that drivers would comply with spatial yielding were 1.079 and 3.612 times for pedestrians who crossed between the stripes and yield line and between the yield line and APCSs, respectively. This was because pedestrians were already beyond the yield line; therefore, if a driver yielded right of way, the chance was very high that the yielding would occur before the yield line. On the pedestrian activities side, drivers were more likely to comply with the spatial yield if pedestrians were crossing while holding or carrying things (OR=1.353), and when pedestrians were using their phones for text or calling (OR=5.89). On the other hand, the odds for drivers to comply with spatial yield were 0.556 and 0.824 when pedestrians were crossing while pushing things and biking, respectively. The most probable reason, especially for bikers, was that they were likely to jaywalk.

The time of the day was also associated with spatial yielding compliance. Drivers were statistically significantly less likely to comply with spatial yielding between 11:00am and 1:00pm (OR=0.575) and between 4:00pm and 6:00pm (OR=0.510). The possible reason for this was that during this time, most drivers were rushing either for lunch or to go home; therefore, they did not concentrate much on pedestrians in the crosswalks.

5.2.Models for pedestrian-infrastructure interactions

In this section, the developed binary models present the association between various dependent variables and covariates. They show the influence of covariates in the probability of occurrence of an event of interest. The events of interest for this case are pedestrian compliance to use the pushbutton, as well as spatial and temporal crossing compliance.

5.2.1. Pedestrians pushbutton compliance

The descriptive analysis revealed that on average about 54% to 66% of pedestrians pressed button to request utilizing the crosswalk, depending on the arrival sequence. The utilization of the pushbutton varied with the signal type and other existing conditions at the crosswalks. The logistic regression associated the pushbutton compliance to other covariates. The results are subdivided into two sections: the first (Table 24) presents the logistic regression results using all data; while the second (Table 24) presents the results for flash-based signal types only. This subdivision was performed to observe the changes in the coefficients and significance levels of the variables, since, it is assumed that the type of signal may play great role for pedestrians' decisions to press the pushbutton.

Crosswalk characteristics play a great role in pedestrians pushing the button before crossing. Table 24 shows that when all signal types are considered, pedestrians crossing at the crosswalks with CFBs and CRFBs signals were about 4.5 times more likely to push the button, compared to those who crossed at TCSs and PHBs. Additionally, pedestrian crossing at RRFBs signals resulted in the odds of 3.6 of pushing the button. The high odds of pushing the buttons can be explained by the fact that TCSs and PHBs are designed similarly to the normal traffic signals at the intersections; thus, pedestrians have a notion that regardless if they press pushbutton or not, they will get the "WALK" phase eventually.

Furthermore, both residential and commercial land use have shown positive associations with pedestrian pushing the button. The odds for pedestrians pushing the button at a crosswalk located in residential and commercial settings were about 2.9 and 1.3, respectively. The more lanes available at the crosswalk, the higher the chance that pedestrians would press the button.

Table 24. LR results for pushbutton compliance for all signal types

Pr(Push button=1)	Coef	OR	Std. Error	z-statistic	P-value
Crosswalk characteristics					
Signal type					
CFBs & CRFBs	1.513	4.541	0.247	6.114	0.000
RRFBs	1.283	3.606	0.207	6.192	0.000
Land use					
Residential	1.065	2.901	0.179	5.948	0.000
Commercial	0.274	1.316	0.181	1.513	0.130
Number of lanes					
Six lanes	0.253	1.287	0.253	0.997	0.319
Seven lanes	-0.112	0.894	0.226	-0.497	0.619
Eight to ten lanes	0.808	2.243	0.308	2.624	0.009
Pedestrian fatal crashes (2013-2016)					
One	-1.098	0.334	0.393	-2.797	0.005
Two	-0.717	0.488	0.349	-2.054	0.040
Active flashes	-1.061	0.346	0.161	-6.581	0.000
Traffic characteristics					
Vehicles within ECD					
Few (five to ten)	0.579	1.784	0.189	3.064	0.002
Platoon (more than ten)	0.581	1.788	0.215	2.709	0.007
Incoming vehicle speed (mph)					
Less than 35	0.133	1.142	0.221	0.602	0.547
Between 35 and 45	0.512	1.668	0.185	2.772	0.006
Greater than 45	0.643	1.903	0.186	3.462	0.001
Vehicle’s position from marked stripes					
Within 40ft	-0.022	0.978	0.226	-0.097	0.922
Between 40ft and 80ft	0.985	2.678	0.338	2.916	0.004
Beyond 80ft	0.117	1.124	0.156	0.751	0.453
Pedestrian characteristics					
Pedestrians activities before crossing					
Holding/carrying stuffs	-0.560	0.571	0.192	-2.924	0.003
Pushing stuffs (bag, stroller, cart)	-0.142	0.868	0.261	-0.542	0.588
Riding bike	-0.708	0.492	0.198	-3.572	0.000
On phone	-0.316	0.729	0.367	-0.861	0.389
Pedestrians' age					
Children and teens	-1.256	0.285	0.272	-4.619	0.000
Young adult	-0.621	0.537	0.238	-2.610	0.009
Adult	-0.824	0.438	0.214	-3.848	0.000
Elderly	-1.141	0.320	0.332	-3.433	0.001
Pedestrian crossing zone					
Between stripes and yield line	-3.025	0.049	0.327	-9.261	0.000
Male pedestrian	-0.289	0.749	0.122	-2.363	0.018
First to arrive	2.061	7.853	0.151	13.687	0.000
Constant	-2.104	0.122	0.460	-4.577	0.000
	Model parameters				
			Number of observations = 2407		
			AIC = 2158		

Table 24 shows that for crosswalks with eight or more lanes, the odds of pressing the button were 2.2 times that of when the crosswalk had only five lanes. Six and seven lane roadways had statistically insignificantly odds. Crash history is negatively associated with pushbutton pressing. Locations with higher numbers of pedestrian involved fatal crashes were found to have negative associations to button pressing. In fact, pedestrians crossing at locations with one and two fatal crashes within a three-year period (2013-2016) were 66.4% and 51.2%, respectively, less likely to press the pushbutton before crossing. The crosswalk flash status the moment a pedestrian arrived at the crosswalk had influence on pushing button. If the flashes were active, pedestrians were 65.4% less likely to press the pushbutton.

Pedestrian-related factors also influenced compliance in pushing a button to cross. Pedestrian age and gender, as well as the crossing zone, sequence of arrival, and activities before crossing were associated with pedestrians pushing the button. Related to age, mixed age was compared to other ages. The results showed that when pedestrians with mixed ages were crossing, the chance that they would press a button was higher than when a specific age was crossing. This can be revealed by the low odds ratios for children and teens (OR = 0.285), young adults (OR = 0.537), adults (OR = 0.438), and the elderly (OR = 0.320). The elderly group has the lowest odds ratio, followed by children and teens. Another important factor is the crossing zone; the odds of pushing the button for pedestrians who crossed between the marked stripes and yield line were about 0.049 compared to those who crossed within marked stripes. Pedestrians' activities before crossing revealed that those who were riding bikes had the lowest odds of pressing the button compared to pedestrians walking normally. Another activity that was statistically significant was holding/carrying things. Pedestrians who held or carried things had 0.571 odds of pressing the pushbutton. Other activities, such as pushing a stroller/bag and using a mobile phone were found not to be statistically significant, at a 95% level. The arrival sequence at the crosswalk revealed that pedestrians who were first to arrive, were about eight times more likely to press the pushbutton. The final pedestrian related factor is the gender of the pedestrian. Male pedestrians were 25.1% less likely to press the button as compared to females.

Vehicular traffic-related factors also play a great role in influencing the use of a pushbutton before crossing. The assessment on this aspect was based on the number, speed, and position of the vehicles at the moment pedestrians arrived. As was expected, higher speeds of vehicles positively influenced the pushbutton usage. Compared to when there was no vehicle coming or traffic had stopped, the speed of more than 45mph was found to nearly double the odds of pedestrians pushing the button. The speeds between 35 and 45 mph was associated with 1.668 odds of pressing the button, while there was no statistically significant difference for speeds lower than 35mph. Number of vehicles within the ECD showed that, the more vehicles, the higher the odds for pressing the button. The base category for this variable was “less than five vehicles,” for which, when compared to five to ten vehicles and ten or more vehicles within ECDs, the odds were 1.784 and 1.788 respectively. The final variable studied is the leading vehicles’ position the moment a pedestrian arrived at the crosswalk. Results in Table 24 show that there was a mixed association between the position of the leading vehicles and pushbutton pressing tendency; however, only the distance between 40 and 80 ft was statistically significant, at a 95% level. At that distance, pedestrians were 2.678 times more likely to press the button, as compared to the when there were no vehicles.

5.2.2. Pedestrians’ spatial crossing compliance

The models for spatial crossing compliance evaluate the pedestrian-infrastructure interaction, by which the compliance of pedestrians to designated crossing zones is assessed. For this case, since according to the law, drivers should yield to pedestrians within the crosswalk, only marked stripes are considered the designated crossing zone. Thus, for a pedestrian to be considered as a spatial-crossing compliant, he/she should start and finish the crossing within the marked stripes. The logistic regression results for spatial crossing compliance are presented in Table 25.

According to the model results in Table 25, pedestrians are less likely to comply at RRFBs as well as at CFBs and CRFBs compared with TCSs and PHBs. The odds ratio for RRFBs was found to be 0.559, while that of CFBs and CRFBs was 0.938. However, the CFBs and CRFBs showed no statistically significant difference from TCSs and PHBs, at a 95% confidence level. Both residential and commercial land use areas are associated with high spatial crossing compliance. At commercial locations, pedestrians were about four times more likely to comply, while the odds were two times in the residential locations, compared with mixed land use areas.

The cross stages at the crosswalks were also found to associate with spatial yielding compliance. Pedestrians were nearly 2.5 times more likely to comply when using a crosswalk that has a single stage than the one with two stages. The optional one/two stages had nearly two times more compliance compared to two stages. The number of road lanes where a crosswalk is located also has significant influence on spatial crossing compliance. The greater number of lanes available, the higher the compliance, as can be revealed by the high odds 2.093, 1.255, and 3.203 for six, seven, and eight to ten lanes, respectively, as compared to a four-lane roadway.

The vehicular crash history of the crosswalk provided inconclusive results, as the increased number of crashes revealed positive and negative influences on crossing compliance. That is, locations that had 10 to 20 crashes (between 2013 and 2016) were associated with less yielding compliance, with and odds ratio of 0.47, while locations that had more than 20 crashes had increased yielding compliance by 39.9%.

Table 25. LR results for spatial crossing compliance for all signal types

Pr(Cross within marked stripes=1)	**Coef**	**OR**	**Std. Error**	**z-statistic**	**P-value**
Crosswalk characteristics					
Signal type					
CFBs & CRFBs	-0.064	0.938	0.361	-0.177	0.860
RRFBs	-0.582	0.559	0.344	-1.694	0.090
Land use					
Residential	0.814	2.258	0.217	3.759	0.000
Commercial	1.427	4.164	0.205	6.974	0.000
Cross stages					
Optional one/two	0.660	1.934	0.354	1.864	0.062
Strictly one	0.909	2.482	0.330	2.755	0.006
Number of lanes					
Six lanes	0.739	2.093	0.265	2.783	0.005
Seven lanes	0.227	1.255	0.348	0.651	0.515
Eight to ten lanes	1.164	3.203	0.346	3.359	0.001
Number of crashes (2013-2016)					
Between 10 and 20	-0.755	0.470	0.275	-2.748	0.006
Greater than 20	0.336	1.399	0.276	1.216	0.224
Use crosswalk sign	0.826	2.285	0.316	2.614	0.009
Active flashes	-0.350	0.705	0.170	-2.060	0.039
Traffic characteristics					
Vehicles within ECD					
Few (five to ten)	0.450	1.568	0.179	2.514	0.012
Platoon (ten or more)	0.825	2.283	0.209	3.948	0.000
Incoming vehicle speed (mph)					
Less than 35	0.172	1.187	0.216	0.793	0.428
Between 35 and 45	0.230	1.258	0.186	1.236	0.216
Greater than 45	0.597	1.816	0.187	3.187	0.001
Vehicle’s position from marked stripes					
Within 40ft	-0.485	0.616	0.263	-1.843	0.065
Between 40ft and 80ft	0.783	2.188	0.366	2.142	0.032
Beyond 80ft	-0.013	0.987	0.157	-0.083	0.934
Pedestrian characteristics					
Gender of pedestrian(s)					
Female(s)	0.138	1.148	0.204	0.675	0.499
Male(s)	-0.309	0.735	0.182	-1.695	0.090
Pedestrians' age					
Children and teens	-0.845	0.429	0.272	-3.103	0.002
Young adult	-0.382	0.683	0.251	-1.522	0.128
Adult	0.163	1.176	0.233	0.697	0.486
Elderly	-0.382	0.683	0.362	-1.055	0.292
Pedestrians activities when crossing					
Holding/carrying stuffs	0.236	1.266	0.214	1.104	0.270
Pushing stuffs (bag, stroller, cart)	1.174	3.235	0.414	2.835	0.005
Riding bike	-0.856	0.425	0.139	-6.167	0.000
On phone	0.221	1.247	0.575	0.384	0.701
First to arrive	0.543	1.720	0.160	3.385	0.001
Pedestrian to/from the bus	-1.154	0.315	0.170	-6.797	0.000
Temporal factors					
Time of the day					
Between 11:00am and 1:00pm	-0.286	0.751	0.229	-1.247	0.212
Between 1:00pm and 4:00pm	-0.247	0.781	0.184	-1.344	0.179
Between 4:00pm and 6:00pm	-0.252	0.777	0.181	-1.391	0.164
Past 6:00pm	-0.456	0.634	0.233	-1.958	0.050
Constant	-0.169	0.845	0.536	-0.315	0.753

Model parameters

Number of observations = 2638

AIC = 2156

The “USE CROSSWALK” signs at the crosswalks were also found to be associated with high crossing compliance. Results in Table 25 show that pedestrians are 2.3 times more likely to comply crossing within marked stripes in the presence of the “USE CROSSWALK” signs. The final crosswalk related factor that affect spatial crossing compliance is the status of the flashing lights or walk signal when a pedestrian arrived at the crosswalk. The probability that a pedestrian would comply with spatial crossing declined by about 30% when he/she arrived at the crosswalk while either flashes were active for RRFBs, CFBs, and CRFBs or the walk signal was on for TCSs and PHBs.

Several personal factors were also assessed in connection to spatial crossing compliance. Among the factors that were found statistically significant affecting spatial yielding compliance include gender and age, activities when crossing, and whether pedestrian is coming or going to the bus. Staring with gender, males were 36.5% less likely and females 14.8% more likely to comply, as compared to mixed gender (male and female) pedestrians. However, the female only category was not statistically significant, at a 95% level. According to Table 25, children, teens, young adults, and the elderly are the age groups that are less likely to comply to spatial crossing when they are using the crosswalk. The results show that children and teens were 31.7% and young adults and the elderly were 31.7% are less likely to comply. Meanwhile, only the children and teen-age groups were statistically significantly different from mixed ages in spatial crossing compliance, at a 95% level.

Referring to the activities when crossing, all of the activities, except riding a bike, are associated with high compliance. To be specific, pedestrians who crossed while holding/carrying things had an odds ratio of 1.266, those who were pushing things (bag, stroller, and cart) had an odds ratio of 3.235, while those who were on the phone had an odds ratio of 1.247. On the other hand, the bike riders had an odds ratio of 0.4.25, as compared to those who crossed normally. The use of a phone and holding/carrying things were not statistically significant, at a 95% level. Pedestrians either originating from or destined towards the bus were found to have a low odds ratio of complying spatially when crossing. This category of pedestrians was 68.5% less likely to comply with spatial crossing. Further, the arrival sequence has shown a great influence in crossing compliance. That is, pedestrians who were first to arrive at the crosswalk were 1.7 times more likely to comply, as compared to the follow up pedestrians.

The existing traffic conditions at a crosswalk also play an important role in the spatial crossing compliance of pedestrians. The traffic conditions covered in this model include number of vehicles within ECDs, incoming vehicle speed (mph), and front vehicle's position from marked stripes when pedestrian arrived at the crosswalk. The number of vehicles is associated with increased compliance, as the more the vehicles the higher the compliance. This was revealed by the 1.568 and 2.283 odds ratios of spatial yielding compliance in the presence of medium (five to ten vehicles) and platoon (more than ten) of vehicles, respectively. The same is true for the incoming vehicle speeds: the higher the speeds, the higher the compliance. In this case, pedestrians were about 26% more likely to comply when the incoming vehicles' travel speeds were between 35 and 45 mph; they were also about two times likely to comply when the speed was greater than 45 mph. The position of the vehicle when the pedestrian arrived at the crosswalk showed a mixed result. Table 25 shows that only when vehicles were within 40 and 80 feet, pedestrians were about two times more like to comply. On the other hand, when vehicles were either too close (within 40 feet) or too far (beyond 80 feet), pedestrians were less likely to comply.

Considering the time of the day, compared to the morning time, pedestrians were less likely to comply for the rest of the day. However, only night time (past 6:00 pm) was statistically significant, at a 95% level. These results show that pedestrians were 24.9%, 21.9%, 22.3%, and 36.6% less likely to comply during the afternoon from 11:00 am-1:00pm, 1:00pm-4:00pm, and 4:00pm-6:00pm, as well as past 6:00pm, respectively.

5.2.3. Pedestrians temporal crossing compliance

Temporal crossing compliance analysis focused on determining the factors associated with compliance to waiting for the walk signal for pedestrians who want to use a crosswalk. In this analysis, only data collected from TCSs and PHBs signals were used, since for flashers, pedestrians normally did not wait for the walk signal. As it has been presented previously, variables were divided into crosswalk-based, pedestrian-based, and traffic-based.

According to the results presented in Table 26, land use, cross stages, and time to yield are the three crosswalk related variables that were found to statistically significantly associated with temporal yielding compliance. Both residential and commercial land uses are associated with less likely to temporal crossing compliance, as compared to mixed land use. In fact, pedestrians in commercial areas presented the worst scenario in terms of temporal crossing compliance, as shown by the odds ratio of 0.078, which was significantly lower than that of residential land use (0.160). The cross stages also revealed that the crosswalks that have one cross stage are associated with higher temporal crossing compliance (OR = 3.753) as compared to the two cross-stage crosswalks. On the other hand, the crosswalks where pedestrians could choose to have either one or two stages of crossing were associated with low temporal crossing compliance (OR = 0.411). The crosswalks by which vehicles took long time to yield were associated with low temporal crossing compliance. In fact, according to Table 26, for every second that pedestrians waited for vehicles to yield, the chance that they might violate temporal crossing compliance increased by 97.8%.

Table 26. LR results for temporal crossing compliance

Pr(Temporal crossing compliance=1)	Coef	OR	Std. Error	z-statistic	P-value
Crosswalk characteristics					
Land use					
Residential	-1.831	0.160	0.419	-4.372	0.000
Commercial	-2.557	0.078	0.405	-6.319	0.000
Cross stages					
Optional one/two	-0.890	0.411	0.356	-2.498	0.012
Strictly one	1.322	3.753	0.446	2.964	0.003
Traffic characteristics					
Vehicles within ECD					
Medium (five to ten)	0.861	2.365	0.306	2.813	0.005
Platoon (ten or more)	0.768	2.155	0.329	2.331	0.020
Time to yield	-0.023	0.978	0.002	-9.897	0.000
Pedestrian characteristics					
Pedestrian crossing zone1					
Between stripes and yield line	-0.392	0.676	0.616	-0.636	0.524
Between yield line and APCS	-1.620	0.198	0.440	-3.681	0.000
Gender of pedestrian(s)					
Female(s) only	0.292	1.339	0.329	0.887	0.375
Male(s) only	-1.106	0.331	0.303	-3.651	0.000
Pedestrians' age					
Children and teens	-0.876	0.416	0.459	-1.910	0.056
Young adult	-0.121	0.886	0.400	-0.303	0.762
Adult	-0.328	0.720	0.343	-0.957	0.339
Elderly	0.111	1.117	0.589	0.188	0.851
Constant	5.229	186.544	0.684	7.647	0.000

Model parameters

Number of observations = 864

AIC = 621.27

Pedestrian-related factors for temporal crossing compliance presented in Table 26 are age, gender, and crossing zone. According to age, all ages except the elderly were associated with less compliance of temporal yielding. However, only children and teens were statistically significant, at a 90% level. Results showed that being a kid or teen, a pedestrian is 58.4% less likely to wait for the walk signal. The gender of the pedestrians revealed that males were about 67% less likely to comply with temporal crossing, while females were about 34% more likely to comply, but females were not statistically significant, at a 90% level. According to the crossing zone, the farther the pedestrian crossed away from the marked stripes, the lower the compliance. This means, pedestrians who crossed between the stripes and yield line were 32.4% less likely to comply, while those who crossed between the yield line and APCS were 81.2% less likely to comply.

Number of vehicles within the ECDs is the only traffic related factor that was found to be associated with the temporal crossing compliance. As was expected, the more the vehicles within ECD, the higher the temporal crossing compliance. The crossing compliance more than doubled when there were five or more vehicles in the ECDs, as compared to when there were less than five vehicles.

5.2.4. Pedestrians spatiotemporal crossing compliance

In a real sense, for crosswalks with TCSs and PHBs, both spatial and temporal crossing compliance occur jointly. Pedestrians may comply with either one, both, or none of the compliances. Due to this situation this study proposed three possible models; multinomial logit, ordered logit, and logistic regression to model the spatiotemporal crossing compliance. The performances of the three competing models were compared and the results are presented in Figure 27.

It can be observed that logistic regression outperformed the multinomial logit (MNL) and ordered logit (Ologit) methods. This is evident based on the BIC scores, AIC scores, and misclassification error rates. The MNL has the highest BIC score (693), followed by Ologit model, while the Logistic regression has the lowest BIC score (496). Similary, the AIC score is relatively high for MNL and Ologit, compared to that of logistic regression. The BIC score difference between the logistic model and Ologit was 117, while the AIC score difference between logistic model and MNL was 105. These differences are significantly higher than the minimum cutoff difference for deciding on the best model. Moreover, the miscalssification error rates followed a similar trend as the AIC score.

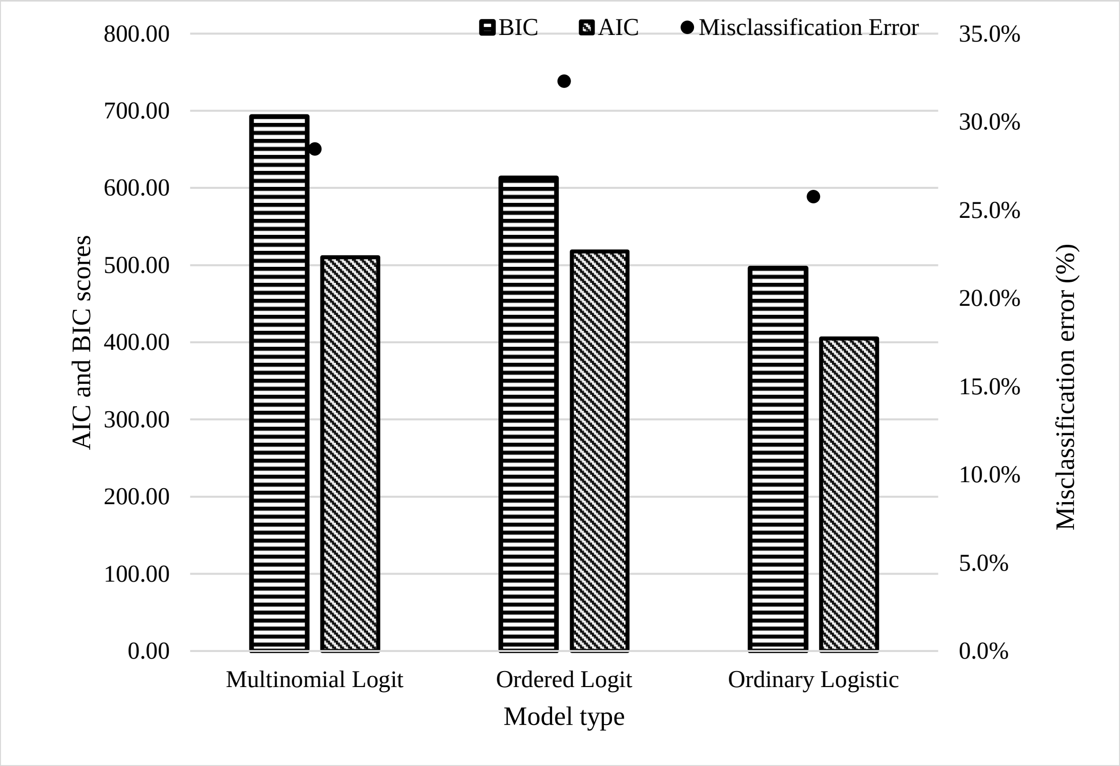


Figure 27. Model performance results

The error was largest (32.3%) for Ologit model and smallest (25.8%) for the logistic model. The difference in misclassification error rates between the logistic model (smallest) and the MNL (intermediate) was 2.7%. Based on these model performance results, it could be concluded that logistic regression was the best model among the three competing models. The following section avails the results discussion based on the logistic regression results.

Table 27. Model results for spatiotemporal crossing compliance

	Ordered logit			Multinomial Logit			Logistic regression		
Variables	**Odds Ratio**	**Std. Err.**	**P-value**	**Odds Ratio**	**Std. Err.**	**P-value**	**Odds Ratio**	**Std. Err.**	**P-value**
Crosswalk characteristics									
Land use									
Residential	0.340	0.377	0.004	0.209	0.427	0.000	0.250	0.407	0.001
Commercial	0.596	0.537	0.335	0.468	0.587	0.196	0.546	0.565	0.284
Cross stages									
Optional one/two	0.214	0.509	0.002	0.261	0.542	0.013	0.226	0.532	0.005
Strictly one	1.193	0.506	0.728	1.495	0.533	0.451	1.372	0.524	0.546
Active WALK signal	1.752	0.204	0.006	1.282	0.224	0.269	1.541	0.217	0.046
Pedestrians' delay	0.989	0.002	0.000	0.987	0.002	0.000	0.986	0.002	0.000
Traffic characteristics									
Vehicle's position from stripes									
Within 40 ft	0.539	0.407	0.128	0.631	0.438	0.294	0.568	0.430	0.188
Between 40ft and 80ft	1.498	0.913	0.658	1.208	0.912	0.836	1.316	0.932	0.769
Beyond 80 ft	0.243	0.325	0.000	0.255	0.360	0.000	0.227	0.349	0.000
Incoming vehicle speed									
Less than 35	0.608	0.454	0.272	0.802	0.500	0.658	0.800	0.476	0.640
Between 35 and 45	0.240	0.280	0.000	0.240	0.304	0.000	0.223	0.298	0.000
Greater than 45	1.068	0.307	0.830	1.462	0.337	0.260	1.319	0.328	0.399
Pedestrians characteristics									
Gender of pedestrian(s)									
Female(s) only	1.315	0.292	0.348	2.361	0.320	0.007	1.890	0.310	0.040
Male(s) only	0.576	0.274	0.044	0.683	0.292	0.192	0.612	0.287	0.087
Pedestrians' age									
Children and teens only	0.422	0.444	0.052	0.284	0.498	0.012	0.320	0.481	0.018
Young adults only	1.162	0.358	0.675	0.921	0.392	0.834	0.992	0.372	0.983
Adults only	0.950	0.309	0.868	0.657	0.330	0.203	0.757	0.319	0.383
Elderly only	0.939	0.551	0.909	0.616	0.589	0.410	0.700	0.582	0.541
Secondary activity involvement									
Holding/carrying stuffs	1.486	0.315	0.210	1.268	0.341	0.486	1.383	0.334	0.331
Pushing stuffs (bag, stroller, cart)	0.996	0.487	0.993	0.793	0.538	0.666	0.875	0.529	0.801
Riding bike	0.438	0.305	0.007	0.471	0.350	0.032	0.451	0.341	0.020
On phone	0.309	0.687	0.087	0.382	0.779	0.217	0.306	0.713	0.097
Constant				66.879	0.877	0.000	60.860	0.854	0.000
/cut1	-7.004	0.876							
/cut2	-3.804	0.817							
Model fit parameters	Pseudo R2 = 0.2109			Pseudo R2 = 0.2822			Pseudo R2 = 0.2517		
	Log likelihood = -406.98			Log likelihood = -370.19			Log likelihood = -318.67		
	LR chi2(22) = 217.53			LR chi2(44) = 291.11			LR chi2(22) = 214.39		
	Prob > chi2 = 0.000			Prob > chi2 = 0.000			Prob > chi2 = 0.000		

Table 27 presents the model results of spatiotemporal crossing compliance for the three competing models. However, since the logistic regression model was found to be the best model, the discussion below focuses on the logistic regression results only. The MNL and Ologit results are briefly discussed in comparison to the logistic regression results. The logistic regression's results discussion is divided into three sections according to the variables category which are: crosswalk-related, pedestrian-related, and traffic-related factors. For the MNL, the base category was the partial-compliance, and only the results for full compliance are presented in Table 27.

The model comparison is based on the changes in p-values and coefficients/odds ratios. In terms of p-values, not much change was observed. Only gender variables have shown much change in p-values. In this variable, the group with female(s) only crossing is statistically significantly different from the group with mixed gender (males and females), at a 95% confidence level for MNL and logistic regression, but statistically insignificant for Ologit. Meanwhile, males are statistically insignificant for MNL. Other variables/categories whose p-values significantly changed include active WALK signal, children and teens, and use of phones. Focusing on the changes in the coefficients/odds ratios, in most cases the odds ratios for logistic regression appear to be the lowest. There was no case that the logistic regression appears to have the largest odds ratios.

According to the logistic regression results Table 27, both residential and commercial land uses are associated with low spatiotemporal crossing compliance. For crosswalks located in commercial land use, pedestrians were about 45% (OR = 0.546) less likely to comply, while the odds were 0.250 that pedestrians in residential locations would comply, compared with pedestrians in mixed land use areas. However, the commercial land use was not a statistically significant variable, at a 95% confidence level. These results were similar to those reported by (K. Kim, Made, and Yamashita 2008), who observed that residential land use was associated with higher violations of spatial crossing compliance in Hawaii.

The cross stages at the crosswalks were also found to associate with spatiotemporal crossing compliance. Pedestrians were about 37% more likely to comply when crossing at a single staged crosswalk. On the other hand, the spatiotemporal crossing compliance declined by about 77% if pedestrians crossed at the optional one/two stages crosswalk. Similar results were reported by (Koh and Wong 2014). The pedestrians' delay is negatively associated with the pedestrians' spatiotemporal crossing compliance. That is, for every second that a pedestrian waits for the crossing phase, the chance that he/she complies with spatiotemporal crossing compliance decreases by 1.4%. The waiting time was also reported by (Brosseau et al. 2013) as one of the factors that were associated with the violation of temporal crossing compliance. The last crosswalk related factor that affect spatiotemporal crossing compliance is the status of the walk signal when a pedestrian arrived at the crosswalk. It was observed that the probability that a pedestrian would comply with spatiotemporal crossing rose by about 1.5 times when he/she arrived at the crosswalk while the walk signal was active.

Several personal factors were also assessed in connection to the spatiotemporal crossing compliance. Among the factors that were found statistically significant affecting spatial yielding compliance include gender, age, and secondary activities when crossing. Starting with gender, when crossing involves individual(s) of one gender alone, male(s) were about 38% less likely, while female(s) were 89% more likely to comply, as compared to when the crossing involved mixed (male(s) and female(s)) pedestrians. Two of the previous studies (K. Kim, Made, and Yamashita 2008; Zhou et al. 2013) found similar results regarding males violation of spatial crossing compliance. According to Table 27, only children and the teen-age group were statistically significantly different from mixed ages in spatial crossing compliance, at a 95% level. The odds of spatiotemporal crossing compliance for this group of individuals was 68% lower compared to when the mixed group was crossing. This finding is contrary to the results reported by (K. Kim, Made, and Yamashita 2008). Referring to the activities when crossing, all the activities, except for holding things were associated with low compliance. To be specific, pedestrians who crossed while holding/carrying things had an odds ratio of 1.383, those who were pushing things (bag, stroller, and cart) had an odds ratio of 0.875, while those who were on the phone had an odds ratio of 0.306. On the other hand, the bike riders had an odds ratio of 0.451, as compared to those who crossed normally. The pushing things (bag, stroller, and cart) and holding/carrying things groups were not statistically significant, even at a 90% level.

The existing traffic condition at the crosswalk also plays an important role in spatiotemporal crossing compliance of pedestrians. The traffic conditions covered in this model include incoming vehicle speed (mph), and front vehicle's position from the marked stripes when a pedestrian arrived at the crosswalk. The incoming vehicle speeds were associated with the increased compliance, as the higher the speeds, the higher the compliance. In this case, pedestrians were a more than 30% (OR=1.319) likely to comply if the incoming vehicle speed was greater than 45mph, but this category was not statistically significant, at a 95% level. On the other hand, when speeds were low, i.e. less than 35mph and between 35 and 45 mph, pedestrians were likely to comply. The position of the front vehicle was also associated with the pedestrians' spatiotemporal crossing compliance; however, the association is only statistically significant for vehicles that are more than 80ft away. According to the results in Table 27, the odds for pedestrians' spatiotemporal crossing compliance were lower when the incoming vehicle was far from the crosswalk at the moment the pedestrians arrived. In fact, if pedestrians arrived at the crosswalk while the nearest vehicle was more than 80ft away, the odds of spatiotemporal crossing compliance declined by 77%. For the cases when the vehicles were between 40ft and 80ft, the odds increased by about 32%, while when the vehicle were within 40ft, the odds declined by 0.2% and 43%. However, these two categories were not statistically significant, at a 95% confidence level. The results make sense since when pedestrians do not see any incoming vehicle they are more likely to jaywalk (Zhou et al. 2013).

CHAPTER 6: STUDY FINDINGS, RECOMMENDATIONS, AND LIMITATIONS

This chapter presents the summary of findings, recommendations, and study limitations based on the study's objectives and performed analyses. The findings are presented first, followed by recommendations and study limitations. The study's recommendations can be used by city planners and engineers to improve traffic safety at signalized midblock crosswalks. The study limitations can be a starting point for further research for researchers interested in this topic.

6.1.Summary of Findings

The summary of findings is divided into three sections: modeling methodology, driver-pedestrian interactions, and pedestrian-infrastructure interactions. Such a division makes it easy for users to identify their points of interest. For instance, the modeling methodologies is more likely to be consumed by researchers, while the other two sections can be used by policy makers, city planners, and engineers, as well as the public.

6.1.1. Modeling methodologies

Several modeling methodologies were applied in this study to associate the outcome variables and explanatory variables. A number of assumptions and performance measures were used to come to a conclusion on the best model for a particular purpose. Based on the work presented in this study, the following findings on the modeling methodologies can be summarized:

- This study successfully showed the presence of three states of yielding compliance of drivers at signalized crosswalks under interrupted traffic flow. In the presence of three states, this study showed that the yielding compliance can be better modeled using hazard-based models (multistate models) than either binary-based or vehicle count models. The model presented in this study was able to incorporate partial-yield, which had not been considered by previous studies. The binary-based models can only explain whether the event is more likely to occur, but not the time-to-event occurrence. On the other hand, the vehicle counts models do not consider the differences in vehicle speeds and arrangements. The same number of vehicles with different arrangements might spend different duration to clear the same distance.

- In modeling the spatiotemporal crossing compliance of pedestrians at signalized midblock crosswalks, this study developed and compared three models: multinomial logit, ordered logit, and logistic regression. The performance measures used were the Bayesian Information Criterion (BIC), Akaike Information Criterion (AIC), and misclassification error. Based on these performance measures, the logistic regression had the best performance, as it had low AIC and BIC, as well as a low misclassification error.
- Near-miss and partial-yield incidents are rarely observed at signalized midblock crosswalks, as shown in this study. Due to the rarity of the events, both Traditional Logistic Regression (TLR) and Rare Events Logistic Regression (RELR) were applied and compared. It was found that the performances of both TLR and RELR were nearly the same for modeling partial-yield incidents. On the other hand, the RELR performed slightly better than TLR in modeling near-miss events.

6.1.2. Driver-pedestrian interactions

The driver-pedestrian interactions include the yielding compliance of the drivers, as well as the risks associated with the yielding compliance. The findings for driver-pedestrian interactions presented here are derived from the descriptive analysis and modeling results. The findings from descriptive analysis are presented first, followed by the findings from the developed models.

- Through descriptive statistics, this study found that there was a variation of partial-yield and full-yield incidents by signal type. The partial-yield varied between 0% at TCSs to 7.3% at CRFBs, while full-yield varied between 58.5% at CFBs to 89.2% at TCSs. The average durations for traffic flow transitions from non-yield to partial-yield, as well as non-yield to full-yield were significantly longer for PHBs, due to signal design. With respect to flash-based signals, the average duration was longer for RRFBs, followed by CRFBs and CFBs. The maximum states of transition duration were found to be about five minutes for PHBs, four minutes for RRFBs, and about one and a half minutes for CRFBs and CFBs.
- The multistate models result revealed that not all factors that were statistically significantly associated with the transition of traffic flow from non-yield to full-yield were necessarily associated with transitions that involved partial-yield. The factors that had high hazard ratios for all three possible transitions, which implied that they were associated with short durations of transition, included signal types CFB, CRFB, and RRFB, as well as the travel direction when yielding. On the other hand, the common variable with low hazard ratios across all three transitions was the presence of a “YIELD HERE” sign. For specific transitions, high risk ratio factors for non-yield to full-yield included few cross stages, high number of vehicles, high AADT, and pedestrian crossing compliance. The low risk ratio factors included high number of lanes, presence of “State Law” sign, and high number of pedestrians. The non-yield to partial-yield transition took a shorter duration to occur in the presence of many vehicles, while a long duration of the same was expected when the incoming vehicles’ speeds are high, and vehicles were too close when pedestrians arrived at crosswalks. For partial-yield to full-yield transition durations, mixed findings were observed, except for vehicle proximity when pedestrians arrived at the crosswalk and the

number of pedestrians waiting to cross, which were associated with a long duration. Meanwhile, number of lanes and cross stages were found to be associated with both short and long transition durations.

- The probability of a partial-yield occurrence was associated with various crosswalk related factors, traffic characteristics, pedestrian characteristics, and temporal factors. The crosswalk related factors that were found to associate with the increased chances of partial-yield included: CFBs and CRFBs; high number of lanes; both narrow and wide raised median types; and a green traffic signal at the intersection immediate to the crosswalk. Traffic characteristics that were associated with the likelihood of a partial-yield occurrence included number of vehicles within ECD and speeds. Partial-yield were also more likely to occur in the immediate traffic flow direction than in the farther direction. Further, the more pedestrians crossing at one incident, the higher the probability of partial-yield. Focusing on temporal factors, this study revealed that for every second that traffic continued to flow while pedestrians were waiting to cross, the probability of partial-yield occurrence increased by 2.1%.
- This study also found that there was a strong association between partial-yield and near-crash events. The odds ratio was almost five times higher that near-crash events would occur if partial-yield had occurred. Moreover, for every second that vehicles continued flowing while pedestrians were waiting to cross, the chances of near-crash events increased by about three percent. However, both partial-yield and time to yield did not show the highest associations to near-crash events. This study found that near-crash events were highly associated with one cross stage, a high number of lanes, and night time.

- The drivers’ spatial yielding compliance was also investigated. The aim was to associate it with crosswalk signalization, signs, and markings, together with other prevailing conditions, by application of logistic regression. Considering crosswalk features and markings, crosswalks with RRFBs were found to have high odds of spatial yielding compliance; the odds were even higher if the crosswalks were equipped with RRFBs at the APCSs. Both “State Law” and “PED XING” signs were associated with high spatial yielding compliance. Moreover, the distance from marked stripes to the yield line and APCSs was also found to play a great role in spatial yielding compliance. Further, the farther the APCSs from the marked stripes, the higher the spatial yielding compliance, while longer distances from the stripes to the yield lines was associated with decreased spatial yielding compliance. In comparison to mixed land use, commercial and residential land uses were also associated with high spatial yielding compliance. Additionally, the crosswalks with few cross stages were associated with low spatial yielding compliance. Other traffic related factors, such as AADT, pedestrian-related factors, and crossing location/zone, also affected spatial yielding compliance.

6.1.3. Pedestrian-infrastructure interactions

The pedestrian-infrastructure interactions analysis involved the assessment of pedestrians’ behaviors towards the use of crosswalk features before and during crossing. Proper use of the crosswalk features resulted in safer crossing. Apart from crosswalk features, other traffic factors, pedestrian characteristics, and temporal factors were evaluated in connection to pedestrian crossing behavior. Based on the descriptive analysis and models developed in this study, the following conclusions can be made:

- In view of the “LOOK BEFORE CROSSING” sign, this study concludes that its presence has no significant influence on pedestrians’ actions towards looking for incoming vehicles before crossing. The analysis in this study found that regardless of the presence or absence of the “LOOK BEFORE CROSSING” sign, pedestrians looked for incoming vehicles. Therefore, the presence of the sign did not change pedestrians’ behaviors. One of the reasons that may explain the habit of looking for incoming vehicles before crossing is the education that has been provided in most American schools, where children are educated on the procedure to follow before crossing a roadway. People grow-up with this habit, and it becomes a part of their lives.
- Although the law does not explicitly define the marked crosswalk, the area marked by stripes can be considered the marked crosswalk area. In that view, the spatial crossing compliance of pedestrians was evaluated. The analysis of pedestrian spatial crossing compliance found that most pedestrians start and finish their crossings within either the marked stripes or between the marked stripes and yield lines. Only about 7% of pedestrians either started or finalized their crossings in the zone that is between yield lines and advanced pedestrian crossing signs.
- As for directive signs at the crosswalks, initial analysis found that the “USE CROSSWALK” sign has no statistically significant correlation to the spatial crossing compliance of pedestrians. However, this finding was based on the correlation test. Utilizing logistic regressions, factors associated with pedestrian spatial crossing compliance were determined. The presence of a “USE CROSSWALK” sign was found to be associated with increased spatial crossing compliance.

- Other factors found to be associated with spatial crossing compliance are residential and commercial land use, few crossing stages, high number of lanes, high incoming vehicle speed, and a greater number of vehicles within ECD, and arrival sequence. Conversely, time of day other than morning, pedestrians' race, pedestrians going from/to buses, and active flash lights/red signals were associated with low spatial crossing compliance. Other factors including crash prone locations, pedestrians' gender and age, and pedestrian activities during crossing were found to have mixed results.
- The temporal crossing compliance, which is the probability that a pedestrian will wait for a "WALK" sign, was also evaluated. This evaluation was performed for PHB and TCS signals because they are the only signals types that have "WALK" signals. It was found that the number of vehicles within was the only factor associated with increased temporal crossing compliance. On the other hand, a longer time from non-yield to full-yield transition, residential and commercial land use, and pedestrians crossing away from the stripes were associated with low temporal crossing compliance. Pedestrians' genders and ages, as well as few cross stages showed mixed results.
- The third sign that was assessed was "PUSH BUTTON TO TURN ON LIGHTS," which is mostly attached at the pushbuttons. In evaluating the influence of the sign, this study considered both the first arriving and follow-up pedestrians. It was found that there was a strong association between the presence of the sign and pushbutton pressing. The first arriving pedestrians resulted in the strongest association, compared to follow-up pedestrians.
- When considering the rate of pushbutton pressing, the proper approach is analyzing by considering the pedestrians' arrival sequence. The reason for this approach is the fact that,

especially for TCSs and PHBs, most pedestrians who find other people waiting to cross assume that the pushbutton has already been pressed. In this study, separating first arriving and follow up pedestrians showed a significant change in pushbutton pressing rates across all signal types.

- The Logistic Regression models developed to associate pushbutton activation to other covariates found that flash-based signal type, large number of lanes, raised median type, a greater number of vehicles within ECD, high incoming vehicle speed, land use where crosswalk is located, and arrival sequence were associated with high pushbutton pressing compliance. On the other hand, crash history of the crosswalks, pedestrians' activities before crossing, pedestrian ages, pedestrians' genders, and crossing zones were associated with low pushbutton pressing compliance. Being specific, this study found that although flash-based signals have high pushbutton pressing rates, there were variations, whereby RRFBs were found to associate with low pressing rates compared to CFBs and CRFBs. Moreover, male pedestrians alone were found to have lower odds of pressing the button than either females alone or a mixture of males and females. Of all factors presented here, arrival sequence and pedestrian crossing zones showed the strongest associations to pushbutton pressing compliance. Pedestrians who were first to arrive at the crosswalk were about eight times more likely to press the pushbutton, compared to those arrived while there were other people either crossing or waiting to cross. Considering crossing zones, regardless of the zone size, crossing outside of the marked stripes was associated with a decrease in the odds of pressing the pushbutton.
- The logistic regression results for spatiotemporal crossing compliance revealed that a strictly single crossing stage, an incoming speed greater than 45mph, and an active

"WALK" sign are the factors that were positively associated with pedestrians' spatiotemporal crossing compliance. On the other hand, pedestrian waiting time, male pedestrians, and children and teens, as well as people who cross while using a phone or riding a bike were negatively associated with spatiotemporal crossing compliance.

6.2.Recommendations

This part of the dissertation presents the recommendations, which are based on the findings from this study. The focus is on the recommendations that could improve pedestrians' safety at various crosswalks.

- As a significant number of pedestrians were found to not press the pushbutton, this study suggests the use of automatic pedestrian detectors. It is understood that this method has had some shortcomings in the past, such as missed calls and false calls. However, with the improvement in technology, sensors such as LiDAR may be applied at crosswalks in order to detect pedestrians. To avoid false calls, these sensors should be positioned to detect only pedestrians that have stepped their feet into the roadways, not only on the sidewalks. Moreover, the entire area within the yield lines should be covered, as it has been observed that not all pedestrians cross within the marked stripes. The locations to be given special priority include crosswalks located near shopping malls, where pedestrians hold/carry things while crossing the roadways. These automatic detectors should also be considered for locations where teens and children (near schools), as well as the elderly are more likely to cross. Apart from automatic pedestrian detectors, audible devices that constantly remind pedestrians to press the pushbutton may be added in order to raise the pressing frequency/rate.

- Partial-yield events and near-crash events, as well as the duration of states' transitions should be included in the metrics for effectiveness analyses of signalized crosswalks, whenever before and after studies are performed.
- The distance between the marked stripes and yield line should be at least 40ft at all crosswalks. A longer distance between the yield line and marked stripes was found to associate with higher yielding compliance; however, the same factor was also found to be associated with low spatial crossing compliance.
- To improve spatial yielding compliance, this study suggests the utilization of RRFBs or any form of flashing signals, as well as rumbles at advanced pedestrian crossing signs (APCSs). Both the flashing lights and rumbles should be activated either when pedestrians push the button, or when triggered by automatic detectors. In addition, the use of yellow "PED XING" signs at the crosswalks and locating APCSs farther away from the stripes are encouraged. Moreover, more education to pedestrians regarding roadway crossing procedures should be emphasized to deter jaywalking.
- Multistate models should be used in modeling the transition states when drivers are yielding right-of-way to pedestrians at signalized midblock crosswalks, regardless of the presence or absence of partial-yield events. As has been shown and discussed in this study, the yielding of right-of-way to pedestrians is better presented as a function of time. The multistate models depict close to reality scenarios, as compared to binary or negative binomial models.
- This study recommends that the logistic regression to be used for studies that are performed to assess the spatiotemporal crossing compliance of pedestrians at signalized crosswalks.

- For rare events, such as partial-yield and near-miss events at crosswalks, this study recommends the use of Rare Event Logistic Regression (RELR) for modeling.
- This study also recommends speed reductions within crosswalk effective distances. It was observed that crosswalks located on low speed roadways were associated with high yielding compliance of drivers. However, low speeds were also associated with low rates of pedestrians pressing the pushbutton. Therefore, additional treatments need to be considered.

6.3. Study Limitations

This part of the dissertation presents the limitations for this study. These study limitations can be used as the starting points for other researchers who are interested in this topic. The limitations are based on the data collection and analyses.

- The multistate model developed in this study estimates the ratio of time-to-event occurrence in the presence of the variable in question, against the lack of said variable. However, it could be more interesting if the exact time-to-event could be estimated. The current approach does not provide, in detail, the magnitude of time-to-event occurrence given two different options of the same variable, or even different variables.
- The data extraction could be improved using the automatic tracking of pedestrians and drivers. With the evolution of machine learning technology, this can be done by developing algorithms in Python, R, or C++ environments.

REFERENCES

Aalen, Odd, and Soren Johansen. 1978. "An Empirical Transition Matrix for Non-Homogeneous Markov Chains Based on Censored Observations." *Scandinavian Journal of Statistics Scand J Statist* 5(5): 141–50. http://www.jstor.org/stable/4615704 (April 13, 2018).

Akaike, Hirotugu. 1974. "A New Look at the Statistical Model Identification." *IEEE Transactions on Automatic Control* (6). http://bayes.acs.unt.edu:8083/BayesContent/class/Jon/MiscDocs/Akaike_1974.pdf (November 30, 2018).

Al-Kaisy, Ahmed, Guilherme T Miyake, Joey Staszcuk, and Danielle Scharf. 2016. "Motorists' Voluntary Yielding of Right of Way at Uncontrolled Midblock Crosswalks with Rectangular Rapid Flashing Beacons." *Journal of Transportation Safety & Security* 10(4): 303–17.

Andersen, Kragh, · Maja, and Pohar Perme. 2008. "Inference for Outcome Probabilities in Multi-State Models." *Lifetime Data Anal* 14: 405–31. https://link.springer.com/content/pdf/10.1007%2Fs10985-008-9097-x.pdf (November 15, 2017).

Araújo, Artur, Luís Meira-Machado, and Javier Roca-Pardiñas. 2014. "TPmsm : Estimation of the Transition Probabilities in 3-State Models." *Journal of Statistical Software* 62(4): 1–29. http://www.jstatsoft.org/v62/i04/ (December 19, 2017).

Ashur, Suleiman, and Mohammad Alhassan. 2015. *Selection of Pedestrian Crossing Treatments at Controlled and Uncontrolled Locations*. http://dx.doi.org/10.5703/1288284315522 (December 15, 2017).

Beck, Laurie F, Ann M Dellinger, and Mary E O 'neil. 2007. "Motor Vehicle Crash Injury Rates by Mode of Travel, United States: Using Exposure-Based Methods to Quantify Differences." *American Journal of Epidemiology* 166(2).

Bennett, Miles K., Hana Manal, and Ron Van Houten. 2014. "A Comparison of Gateway In-Street Sign Configuration to Other Driver Prompts to Increase Yielding to Pedestrians at Crosswalks." *Journal of Applied Behavior Analysis* 47(1): 3–15.

http://doi.wiley.com/10.1002/jaba.103 (December 28, 2016).

Brewer, Marcus A., Kay Fitzpatrick, and Raul Avelar. 2015. “Rectangular Rapid Flashing Beacons and Pedestrian Hybrid Beacons.” *Transportation Research Record: Journal of the Transportation Research Board* 2519: 1–9. http://trrjournalonline.trb.org/doi/10.3141/2519-01 (December 29, 2016).

Brosseau, Marilyne, Sohail Zangenehpour, Nicolas Saunier, and Luis Miranda-Moreno. 2013. “The Impact of Waiting Time and Other Factors on Dangerous Pedestrian Crossings and Violations at Signalized Intersections: A Case Study in Montreal.” *Transportation Research Part F: Psychology and Behaviour* 21: 159–72. http://dx.doi.org/10.1016/j.trf.2013.09.010 (August 30, 2018).

Carsten, O.M.J, D.J Sherborne, and J.A Rothengatter. 1998. “Intelligent Traffic Signals for Pedestrians: Evaluation of Trials in Three Countries.” *Transportation Research Part C: Emerging Technologies* 6(4): 213–29. https://www.sciencedirect.com/science/article/pii/S0968090X98000163 (March 8, 2018).

Chimba, Deo et al. 2014. “Impact of Abandoned and Disabled Vehicles on Freeway Incident Duration.” *Journal of Transportation Engineering* 140(3). http://ascelibrary.org.ezproxy.library.unlv.edu/doi/pdf/10.1061/%28ASCE%29TE.1943-5436.0000635 (July 28, 2017).

Choirat, Christine et al. 2018. *Package ‘Zelig.’* https://cran.r-project.org/web/packages/Zelig/Zelig.pdf (May 25, 2018).

Cochran, William G. 1977. *Sampling Techniques*. New York: John Wiley and Sons, Inc. https://hwbdocuments.env.nm.gov/Los Alamos National Labs/General/14447.pdf (April 4, 2018).

Datta, Somnath, and Glen A Satten. 2001. “Validity of the Aalen–Johansen Estimators of Stage Occupation Probabilities and Nelson–Aalen Estimators of Integrated Transition Hazards for Non-Markov Models.” *Statistics & Probability Letters* 55: 403–11. https://ac.els-cdn.com/S0167715201001559/1-s2.0-S0167715201001559-main.pdf?_tid=00ac8d66-430a-477b-a6c2-19cb0417316d&acdnat=1523635533_3cb3b93b51eab21125cf37eb339e0aba

(April 13, 2018).

Domarad, J, P Grisak, and J Bolger. 2013. “Improving Crosswalk Safety: Rectangular Rapid-Flashing Beacon (RRFB) Trial in Calgary.” https://trid.trb.org/view.aspx?id=1263306 (December 29, 2016).

Dougald, Lance E. 2015. *Evaluation of a Rectangular Rapid Flashing Beacon System at the Belmont Ridge Road and W&OD Trail Mid-Block Crosswalk*. www. (December 15, 2017).

Eapen, Anna. 2014. “Evaluation of PHB Mid-Street Crossing System in Las Vegas, Nevada - Pedestrian Perspectives.” University of Nevada, Las Vegas . http://digitalscholarship.unlv.edu/thesesdissertations (March 13, 2018).

Fabozzi, Frank J, Sergio M Focardi, Svetlozar T Rachev, and Bala G Arshanapalli. 2014. *The Basics of Financial Econometrics*. https://onlinelibrary.wiley.com/doi/pdf/10.1002/9781118856406.app5 (November 30, 2018).

FHWA. 2008. “Interim Approval for Optional Use of Rectangular Rapid Flashing Beacons (IA-11).” https://mutcd.fhwa.dot.gov/resources/interim_approval/ia11/ia11_rrfb_iapmemo.pdf (November 30, 2017).

———. 2018a. *Interim Approval for the Optional Use of Pedestrian-Actuated Rectangular Rapid-Flashing Beacons at Uncontrolled Marked Crosswalks (IA-21)*. https://static.tti.tamu.edu/tti.tamu.edu/documents/TTI-CTS-0010.pdf (December 28, 2018).

———. 2018b. “Interim Approval IA-11 Termination Memorandum - Interim Approvals.” https://mutcd.fhwa.dot.gov/resources/interim_approval/ia11/terminationmemo/index.htm (April 13, 2018).

Fisher, Donald, and Lisandra Garay-Vega. 2012. “Advance Yield Markings and Drivers’ Performance in Response to Multiple-Threat Scenarios at Mid-Block Crosswalks.” *Accident Analysis and Prevention* 44: 35–41.

Fitzpatrick, Kay, et al. 2016. “Comparison of Above Sign and Below Sign Placement for Rectangular Rapid-Flashing Beacons.” *Transportation Research Record: Journal of the*

Transportation Research Board (2562): 45–52.

Fitzpatrick, Kay et al. 2006. *Improving Pedestrian Safety at Unsignalized Crossings*. https://nacto.org/wp-content/uploads/2010/08/NCHRP-562-Improving-Pedestrian-Safety-at-Unsignalized-Crossings.pdf (March 12, 2018).

———. 2011. *Evaluation of Pedestrian and Bicycle Engineering Countermeasures: Rectangular Rapid-Flashing Beacons, HAWKs, Sharrows, Crosswalk Markings, and the Development of an Evaluation Methods Report*. http://www.fhwa.dot.gov/publications/research/safety/pedbike/11039/11039.pdf (December 29, 2016).

———. 2014. *Characteristics of Texas Pedestrian Crashes and Evaluation of Driver Yielding at Pedestrian Treatments*. http://tti.tamu.edu/documents/0-6702-1.pdf (December 28, 2016).

Fitzpatrick, Kay, Raul Avelar, et al. 2015. *Investigating Improvements to Pedestrian Crossings With an Emphasis on the Rectangular Rapid-Flashing Beacon*. http://www.fhwa.dot.gov/publications/research/safety/15043/15043.pdf (December 29, 2016).

———. 2016. *Evaluation of Pedestrian Hybrid Beacons and Rapid Flashing Beacons*. https://www.fhwa.dot.gov/publications/research/safety/16040/16040.pdf (March 12, 2018).

Fitzpatrick, Kay, Marcus A Brewer, Raul Avelar, and Tomas Lindheimer. 2016. *Will You Stop for Me? Roadway Design and Traffic Control Device Influences on Drivers Yielding to Pedestrians in a Crosswalk with a Rectangular Rapid-Flashing Beacon Final Report*. College Station.

Fitzpatrick, Kay, Marcus Brewer, and Raul Avelar. 2014. "Driver Yielding at Traffic Control Signals, Pedestrian Hybrid Beacons, and Rectangular Rapid-Flashing Beacons in Texas." *Transportation Research Record: Journal of the Transportation Research Board* (2463): 46–54.

Fitzpatrick, Kay, Ingrid B Potts, Marcus A Brewer, and Raul Avelar. 2015. "Comparison of Rectangular and Circular Rapid-Flashing Beacons in an Open-Road Setting." *Transportation Research Record: Journal of the Transportation Research Board* (2492):

69–77.

Foster, Nick, Christopher M Monsere, and Katherine Carlos. 2014. "Evaluating Driver and Pedestrian Behaviors at Enhanced, Multilane, Midblock Pedestrian Crossings Case Study in Portland, Oregon." *Transportation Research Record Journal of the Transportation Research Board Transportation Research Board of the National Academies* (2464): 59–66.

Garay-Vega, Lisandra. 2008. "DRIVERS PERFORMANCE IN RESPONSE TO SIGHT-LIMITED AND MULTI-THREAT CRASH SCENARIOS AT MID-BLOCK CROSSWALKS: AN EVALUATION OF ADVANCE YIELD MARKINGS AND SIGN PROMPTS A Dissertation Presented Civil and Environmental Engineering." University of Massachusetts Amherst.

Gates, Timothy J et al. 2016. *Development of Safety Performance Functions and Other Decision Support Tools to Assess Pedestrian and Bicycle Safety*. https://wmich.edu/sites/default/files/attachments/u883/2017/TRCLC_14_06_Report.pdf (December 15, 2017).

Giard, Nicole, Paul Lichtenstein, and Anatoli I Yashin. 2002. "A Multistate Model for the Genetic Analysis of the Ageing Process." *STATISTICS IN MEDICINE Statist. Med* 21: 2511–26.

Godavarthy, Ranjit Prasad, and Eugene R. Russell. 2016. "Study of Pedestrian Hybrid Beacon's Effectiveness for Motorists at Midblock Pedestrian Crossings." *Journal of Traffic and Transportation Engineering (English Edition)* 3(6): 531–39. http://www.sciencedirect.com/science/article/pii/S2095756415306413 (April 5, 2017).

Guns, M, and V Vanacker. 2012. "Logistic Regression Applied to Natural Hazards: Rare Event Logistic Regression with Replications." *Nat. Hazards Earth Syst. Sci* 12: 1937–47. www.nat-hazards-earth-syst-sci.net/12/1937/2012/ (July 12, 2018).

Guo, Feng, Sheila G Klauer, Michael T McGill, and Thomas A Dingus. 2010. *Evaluating the Relationship Between Near-Crashes and Crashes: Can Near-Crashes Serve as a Surrogate Safety Metric for Crashes?* Blacksburg. www.ntis.gov (July 28, 2018).

Guo, Hongwei et al. 2014. "Modeling the Perceptions and Preferences of Pedestrians on

Crossing Facilities." *Discrete Dynamics in Nature and Society* 2014: 1–8. http://www.hindawi.com/journals/ddns/2014/949475/ (August 30, 2018).

Hamidun, Rizati, Nurul Elma Kordi, Intan Rohani Endut, and Siti Zaharah Ishak. 2016. "Behavioral Observations of Crossing Pedestrians at Urban Signalized Intersection." *Jurnal Teknologi (Sciences & Engineering)* 78(5): 9–14. https://www.researchgate.net/profile/Rizati_Hamidun2/publication/303296028_Behavioral_observations_of_crossing_pedestrians_at_urban_signalized_intersections/links/576f6cd508ae62194746c028.pdf (July 11, 2017).

Hayward, John C. 1972. "Near-Miss Determination Through Use of a Scale of Danger." *Highway Research Record* (384). https://trid.trb.org/view/115323 (July 27, 2018).

Hougaard, Philip. 1999. "Multi-State Models: A Review." *Lifetime Data Analysis* 5: 239–64.

Houten, Ron, J. Malenfant, and Dave McCusker. 2001. "Advance Yield Markings: Reducing Motor Vehicle—Pedestrian Conflicts at Multilane Crosswalks with Uncontrolled Approach." *Transportation Research Record: Journal of the Transportation Research Board* 1773: 69–74. http://trrjournalonline.trb.org/doi/10.3141/1773-08 (July 19, 2018).

Van Houten, Ron et al. 2002. "Advance Yield Markings and Fluorescent Yellow-Green RA 4 Signs at Crosswalks with Uncontrolled Approaches." *Transportation Research Record: Journal of the Transportation Research Board* 1818: 119–24. http://trrjournalonline.trb.org/doi/10.3141/1818-18 (July 19, 2018).

———. 2011. "Pedestrians." In *Handbook of Traffic Psychology*, Kalamazoo, 353–365.

Van Houten, Ron, Ralph Ellis, and Erick Marmolejo. 2008. "Stutter-Flash Light-Emitting-Diode Beacons to Increase Yielding to Pedestrians at Crosswalks." *Transportation Research Record: Journal of the Transportation Research Board* 2073: 69–78. http://trrjournalonline.trb.org/doi/10.3141/2073-08 (December 28, 2016).

Houten, Ron Van, and Louis Malenfant. 2008. *An Analysis of the Efficacy of Rectangular-Shaped Rapid-Flash LED Beacons to Increase Yielding to Pedestrians Using Crosswalks on Multilane Roadways in The*.

Houten, Ron Van, Louis Malenfant, and Ahmos Rolider. 1985. "Increasing Driver Yielding and

Pedestrian Signaling with Prompting, Feedback, and Enforcement." *Journal of Applied Behavior Analysis* 18(2): 1307998. http://www.pubmedcentral.gov/articlerender.fcgi?artid=1307998 (July 27, 2018).

Huang, Herman F, and Charles V Zeegar. 2001. "An Evaluation of Illuminated Pedestrian Push Buttons in Windsor, Ontario." https://www.fhwa.dot.gov/publications/research/safety/pedbike/0102.pdf (April 4, 2017).

Hughes, Ronald, Herman Huang, Charles Zegeer, and Michael Cynecki. 2000. "Automated Detection of Pedestrians in Conjunction with Standard Pedestrian Push Buttons at Signalized Intersections." *Transportation Research Record: Journal of the Transportation Research Board* 1705: 32–39. http://trrjournalonline.trb.org/doi/10.3141/1705-06 (December 30, 2016).

Hunter, William, Raghavan Srinivasan, and Carol Martell. 2012. "Evaluation of Rectangular Rapid Flash Beacon at Pinellas Trail Crossing in Saint Petersburg, Florida." *Transportation Research Record Journal of the Transportation Research Board Transportation Research Board of the National Academies* (2314): 7–13.

Karkee, Ganesh J, Shashi S Nambisan, and Srinivas S Pulugurtha. 2010. "Motorist Actions at a Crosswalk With an In-Pavement Flashing Light System." *Traffic Injury Prevention* 11: 642–49. http://www.tandfonline.com/action/journalInformation?journalCode=gcpi20 (May 17, 2017).

Kidando, Emmanuel, Ren Moses, Eren E. Ozguven, and Thobias Sando. 2017. "Bayesian Nonparametric Model for Estimating Multistate Travel Time Distribution." *Journal of Advanced Transportation* 2017: 1–9. https://www.hindawi.com/journals/jat/2017/5069824/ (April 18, 2019).

Kim, Jiwon, Hani S Mahmassani, and Jing Dong. 2010. "Likelihood and Duration of Flow Breakdown Modeling the Effect of Weather." *Journal of Transportation Research Record*: 19–28. http://trrjournalonline.trb.org/doi/pdf/10.3141/2188-03 (July 16, 2017).

Kim, Karl, Brunner I Made, and Eric Yamashita. 2008. "Modeling Violation of Hawaii's Crosswalk Law." *Accident Analysis and Prevention* 40: 894–904. https://ac.els-cdn.com/S000145750700187X/1-s2.0-S000145750700187X-main.pdf?_tid=2a6b1a23-005b-40fd-a0e0-15f7f9117538&acdnat=1552697252_c09a3e941b48fc713b2bacdc9d74ce46 (March 15, 2019).

King, Gary, and Langche Zeng. 2001. “Logistic Regression in Rare Events Data.” https://gking.harvard.edu/files/0s.pdf (July 10, 2018).

Kneib, Thomas, and Andrea Hennerfeind. 2008. “Bayesian Semiparametric Multi-State Models.” *Statistical Modelling* 72(82): 169–1.

Koh, P P, and Y D Wong. 2014. “Gap Acceptance of Violators at Signalised Pedestrian Crossings.” *Accident Analysis and Prevention* 62: 178–85. http://dx.doi.org/10.1016/j.aap.2013.09.020 (August 30, 2018).

Kuhn, Max. 2017. *Package “Caret.”* https://cran.r-project.org/web/packages/caret/caret.pdf (June 1, 2017).

Kutela, Boniphace, and Hualiang Teng. 2018. “Parameterizing the Yielding Compliance of Motorists at Signalized Midblock Crosswalks Using Mixed Effects Logistic Regression.” In *Transportation Research Board 97th Annual Meeting*, Washington DC, United States.

Laflamme, Eric M, and Paul J Ossenbruggen. 2017. “Effect of Time-of-Day and Day-of-the-Week on Congestion Duration and Breakdown: A Case Study at a Bottleneck in Salem, NH.” *journal of traffic and transportation engineering* 4(1): 3–1. http://dx.doi.org/10.1016/j.jtte.2016.08.004 (July 16, 2017).

Leitgöb, Heinz. 2013. “The Problem of Modeling Rare Events in ML-Based Logistic Regression s Assessing Potential Remedies via MC Simulations.” In *Conference: European Survey Research Association*, https://www.europeansurveyresearch.org/conf/uploads/494/678/167/PresentationLeitg_b.pdf (December 21, 2018).

Levelt, M. 1992. *Improvement of Pedestrian Safety and Comfort at Traffic Lights: Results from French, British and Dutch Field Tests*.

Manual on Uniform Traffic Control Devices (MUTCD). 2009. Washington, DC: FHWA, U.S.

Department of Transportation.

Martinussen, Torben, and Thomas H Scheike. 2006. *Dynamic Regression Models for Survival Data*. New York, NY: Springer New York. http://link.springer.com/10.1007/0-387-33960-4 (April 13, 2018).

Matsui, Yasuhiro, Masahito Hitosugi, Tsutomu Doi, et al. 2013. “Features of Pedestrian Behavior in Car-to-Pedestrian Contact Situations in Near-Miss Incidents in Japan.” *Traffic Injury Prevention* 14(sup1): S58–63.

Matsui, Yasuhiro, Masahito Hitosugi, Kunio Takahashi, and Tsutomu Doi. 2013. “Situations of Car-to-Pedestrian Contact.” *Traffic Injury Prevention* 14(1): 73–77.

Mead, Jill, Charlie Zegeer, and Max Bushell. 2014. *Evaluation of Pedestrian-Related Roadway Measures: A Summary of Available Research Www.Pedbikeinfo.Org*. http://www.pedbikeinfo.org/cms/downloads/PedestrianLitReview_April2014.pdf (December 15, 2017).

Meira-Machado, Luis et al. 2015. “Package ‘P3state.Msm.’” https://cran.r-project.org/web/packages/p3state.msm/p3state.msm.pdf (March 17, 2018).

Meira-Machado, Luís, and Javier Roca-Pardiñas. 2011. “Analyzing Survival Data from an Illness-Death Model.” *Journal of Statistical Software* 38(3): 1–18. http://www.jstatsoft.org/v38/i03/ (March 17, 2018).

Meira-Machado, Luís, Jacobo Uña-Álvarez, Carmen Cadarso-Suárez, and Per K Andersen. 2009. “Multi-State Models for the Analysis of Time-to-Event Data.” *Statistical Methods in Medical Research* 18: 195–222. http://journals.sagepub.com/doi/pdf/10.1177/0962280208092301 (March 15, 2018).

Meira-Machado, Luís, Jacobo de Uña-Álvarez, and Carmen Cadarso-Suárez. 2006. “Nonparametric Estimation of Transition Probabilities in a Non-Markov Illness–death Model.” *Lifetime Data Analysis* 12(3): 325–44. http://link.springer.com/10.1007/s10985-006-9009-x (March 18, 2018).

Mitman, Meghan, Douglas Cooper, and Brooke DuBose. 2010. “Driver and Pedestrian Behavior at Uncontrolled Crosswalks in Tahoe Basin Recreation Area of California.” *Transportation*

Research Record: Journal of the Transportation Research Board 2198: 23–31. http://trrjournalonline.trb.org/doi/10.3141/2198-04 (December 28, 2016).

Mitman, Meghan Fehlig, David R. Ragland, and Charles V. Zegeer. 2008. “Marked-Crosswalk Dilemma.” *Transportation Research Record: Journal of the Transportation Research Board* 2073(1): 86–93. http://journals.sagepub.com/doi/10.3141/2073-10 (July 19, 2018).

Morrissey, Sam. 2013. *Circular Rapid Flashing Beacons*. http://www.dot.ca.gov/trafficops/ctcdc/docs/2013-04-24_Final_rpt_4(09)-8.pdf (March 12, 2018).

Nam, Doohee, and Fred Mannering. 2000. “An Exploratory Hazard-Based Analysis of Highway Incident Duration.” *Transportation Research Part A: Policy and Practice* 34(2): 85–102. http://ac.els-cdn.com/S0965856498000652/1-s2.0-S0965856498000652-main.pdf?_tid=4e29eee4-4ee6-11e7-a133-00000aab0f6b&acdnat=1497213852_c2dc542127b03c0dca849ca5f738b2a0 (June 11, 2017).

Nambisan et al. 2009. “Effectiveness of Automatic Pedestrian Detection Device and Smart Lighting for Pedestrian Safety.” *Transportation Research Record: Journal of the Transportation Research Board*. http://trrjournalonline.trb.org/doi/pdf/10.3141/2140-03 (December 27, 2016).

NDOT. “Be Pedestrian Safe.” http://www.zerofatalitiesnv.com/be-pedestrian-safe/ (December 14, 2017).

NHTSA. 2015. *2015 Data: Pedestrians*.

“NRS: CHAPTER 484B - RULES OF THE ROAD.” https://www.leg.state.nv.us/NRS/NRS-484B.html#NRS484BSec283 (June 1, 2017).

Pécheux, K, J Bauer, and P Mcleod. 2009. *Pedestrian Safety and ITS-Based Countermeasures Program for Reducing*. https://safety.fhwa.dot.gov/ped_bike/tools_solve/ped_scdproj/sys_impact_rpt/sys_impact_rpt.pdf (June 14, 2017).

Porter, Bryan E, Ingrid Neto, Isabel Balk, and Jennika K Jenkins. 2016. “Investigating the

Effects of Rectangular Rapid Flash Beacons on Pedestrian Behavior and Driver Yielding on 25 Mph Streets: A Quasi-Experimental Field Study on a University Campus." *Transportation Research Part F: Traffic Psychology and Behaviour* 42(3): 509–21.

Potts, Ingrid B et al. 2015. "Effect of Beacon Activation and Traffic Volume on Driver Yielding Behavior at Rapid Flashing Beacons." *Transportation Research Record: Journal of the Transportation Research Board* (2492): 78–83.

Prevedouros, Panos D. 2001. "Evaluation of In-Pavement Flashing Lights on a Six-Lane Arterial Pedestrian Crossing." In *ITE 2001 Annual Meeting*, Chicago, IL.

R Core Team. 2018. "R: A Language and Environment for Statistical Computing." https://www.r-project.org/ (November 17, 2018).

Ragland, David, and Meghan Fehlig Mitman. 2007. *Driver/Pedestrian Understanding and Behavior at Marked and Unmarked Crosswalks*. https://escholarship.org/uc/item/1h52s226 (July 19, 2018).

Retting, Richard. 2017. *Pedestrian Traffic Fatalities by State 2017 PRELIMINARY DATA*. www.ghsa.org (April 6, 2018).

Ripley, Brian et al. 2018. *Package "MASS."* http://www.stats.ox.ac.uk/pub/MASS4/ (November 30, 2018).

Ripley, Brian, and William Venables. 2016. *Package "Nnet."* http://www.stats.ox.ac.uk/pub/MASS4/ (November 30, 2018).

Rosenbloom, Tova. 2009. "Crossing at a Red Light: Behaviour of Individuals and Groups." *Transportation Research Part F: Psychology and Behaviour* 12: 389–394. https://ac.els-cdn.com/S1369847809000278/1-s2.0-S1369847809000278-main.pdf?_tid=4e6b807d-efeb-4e5e-94c4-3e15260c93b0&acdnat=1535642904_b48aa5a19459baf0aed9750b250661d0 (August 30, 2018).

Ross, June, Dan Serpico, and Robin Lewis. 2011. *Assessment of Driver Yielding Rates Pre- and Post-RRFB Installation, Bend, Oregon*. Salem, OR.

RTCSN. 2015. *2014 Southern Nevada Household Travel Survey Final Report*. Las Vegas.

https://www.nrel.gov/transportation/secure-transportation-data/assets/pdfs/rtcsnv_2014_final_report.pdf (April 9, 2018).

Samuel, S et al. 2013. "Effect of Advance Yield Markings and Symbolic Signs on Vehicle–Pedestrian Conflicts Field Evaluation." *Transportation Research Record: Journal of the Transportation Research Board* (2393): 139–46. http://journals.sagepub.com/doi/pdf/10.3141/2393-16 (July 19, 2018).

Shurbutt, Jim, and Ron Van Houten. 2010. *Effects of Yellow Rectangular Rapid-Flashing Beacons on Yielding at Multilane Uncontrolled Crosswalks*.

Shurbutt, Jim, Ron Van Houten, Shawn Turner, and Brad Huitema. 2009. "Analysis of Effects of LED Rectangular Rapid-Flash Beacons on Yielding to Pedestrians in Multilane Crosswalks." *Transportation Research Record: Journal of the Transportation Research Board*: 85–95.

Sisiopiku, V.P, and D Akin. 2003. "Pedestrian Behaviors at and Perceptions towards Various Pedestrian Facilities: An Examination Based on Observation and Survey Data." *Transportation Research Part F: Traffic Psychology and Behaviour* 6(4): 249–74. https://www.sciencedirect.com/science/article/pii/S136984780300041X (March 9, 2018).

Spruance, Spotswood L, Julia E Reid, Michael Grace, and Matthew Samore. 2004. "Hazard Ratio in Clinical Trials." *Antimicrobial agents and chemotherapy* 48(8): 2787–92. http://www.ncbi.nlm.nih.gov/pubmed/15273082 (July 11, 2018).

Stone, M. 1979. 41 Source: Journal of the Royal Statistical Society. Series B (Methodological) *Comments on Model Selection Criteria of Akaike and Schwarz*. https://www.jstor.org/stable/pdf/2985044.pdf?refreqid=excelsior%3A11311bb2158e530a23fff27f08c85fc3 (November 30, 2018).

Strategic Highway Safety Plan. 2016. http://www.zerofatalitiesnv.com/wp-content/themes/zero-nv-2015/downloads/2016-fact-sheets/PedestrianSafetyProblem.pdf (January 3, 2017).

Sucha, Matus, Daniel Dostal, and Ralf Risser. 2017. "Pedestrian-Driver Communication and Decision Strategies at Marked Crossings." *Accident Analysis and Prevention* 102: 41–50.

Turner, Fitzpatrick, Brewer, and Park. 2006. "Motorist Yielding to Pedestrians at Unsignalized

Intersections Findings from a National Study on Improving Pedestrian Safety.” *Transportation Research Record Journal of the Transportation Research Board Transportation Research Board of the National Academies*: 1–12.

UCLA: Statistical Consulting Group. 2014. “Introduction to Generalized Linear Mixed Models.” https://stats.idre.ucla.edu/other/mult-pkg/introduction-to-generalized-linear-mixed-models/ (May 31, 2017).

Veazey, Lindsay M et al. 2016. “The Implementation of Rare Events Logistic Regression to Predict the Distribution of Mesophotic Hard Corals across the Main Hawaiian Islands.” *PeerJ*. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4941748/pdf/peerj-04-2189.pdf (July 10, 2018).

Voorhees, Alan M. 2017. *Predictable Is Preventable: Tracking Pedestrian Near-Miss Incidents*. http://www.njcrossingguards.org/wp-content/uploads/2017/11/Ped-Near-Miss-report-5.25.17.pdf (July 27, 2018).

Wang, Yanjun, and Qun Liu. 2006. “Comparison of Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) in Selection of Stock–recruitment Relationships.” *Fisheries Research* 77: 220–25. http://depts.washington.edu/fish437/resources/Week 4/Wang and Liu 2005.pdf (June 4, 2017).

Washington, Simon., Matthew G. Karlaftis, and Fred L. Mannering. 2011. *Statistical and Econometric Methods for Transportation Data Analysis*. CRC Press. https://www.crcpress.com/Statistical-and-Econometric-Methods-for-Transportation-Data-Analysis/Washington-Karlaftis-Mannering/p/book/9781420082852 (November 30, 2018).

Wickham, Hadley, Winston Chang, et al. 2018. *Package ‘Ggplot2.’* https://cran.r-project.org/web/packages/ggplot2/ggplot2.pdf (December 3, 2018).

Wickham, Hadley, Romain François, Lionel Henry, and Kirill Müller. 2018. *Package ‘Dplyr.’* https://cran.r-project.org/web/packages/dplyr/dplyr.pdf (December 3, 2018).

Williams, Richard. 2018. *Analyzing Rare Events with Logistic Regression*. https://www3.nd.edu/~rwilliam/ (December 21, 2018).

Woodridge, Jeffrey M. 2012. *Introductory Economics A Modern Approach*. 5th ed. South-

Western, Cengage Learning. http://economics.ut.ac.ir/documents/3030266/14100645/Jeffrey_M._Wooldridge_Introductory_Econometrics_A_Modern_Approach__2012.pdf (June 9, 2017).

Yanfeng, Wang et al. 2010. “Characteristic Analysis of Pedestrian Violation Crossing Behavior Based on Logistics Model.” In *2010 International Conference on Intelligent Computation Technology and Automation*, IEEE, 926–28. http://ieeexplore.ieee.org/document/5523150/ (August 30, 2018).

Zegeer, Charles V et al. 2005. *Safety Effects of Marked Versus Unmarked Crosswalks at Uncontrolled Locations Final Report and Recommended Guidelines*. https://www.fhwa.dot.gov/publications/research/safety/04100/04100.pdf (December 14, 2017).

Zheng, Yinan et al. 2015. “Modeling Vehicle -Pedestrian Interactions Outside of Crosswalks.” *Simulation Modelling Practice and Theory* 59: 89–101. http://dx.doi.org/10.1016/j.simpat.2015.08.005 (May 15, 2017).

Zhou, Zhu-Ping, Ying-Shun Liu, Wei Wang, and Yong Zhang. 2013. “Multinomial Logit Model of Pedestrian Crossing Behaviors at Signalized Intersections.” *Discrete Dynamics in Nature and Society* 2013: 1–8. https://www.hindawi.com/journals/ddns/2013/172726/ (August 30, 2018).

CURRICULUM VITAE

Boniphace Kutela

Email: kutela4gmail.com

Education:

Ph.D. in Civil Engineering **3.90 GPA**
University of Nevada Las Vegas (UNLV) August 2014 – May 2019

- Dissertation: ***Modeling Driver-Pedestrian-Infrastructure Interactions at Signalized Midblock Crosswalks***

Master of Engineering (Civil) **3.91 GPA**
Tennessee State University (TSU) August 2011 – May 2013

- Thesis: ***The Impacts of Abandoned and Disabled Vehicles on Highway Incidents***

B.S. in Civil and Structural Engineering **3.42 GPA**
University of Dar Es Salaam (UDSM) August 2006 – July 2010

- Final Year Project: ***Analysis of Shell Structures by Finite Element Method***

Work History:

Coordinator for USDOT Railroad University Transportation Center at UNLV	January 2017 to present
Graduate Research and Teaching Assistant, University of Nevada, Las Vegas	August 2014 to present
Assistant Lecturer Ardhi University (Tanzania)	May 2013 – August 2014
Assistant Lecturer St. Joseph University in Tanzania (Part-time)	May 2013 – August 2014
Graduate Research and Teaching Assistant, Tennessee State University	August 2011 – May 2013
Teaching Assistant, Ardhi University (Tanzania)	July 2010–August 2011

Teaching and Mentoring Experience:

Courses taught at University of Nevada, Las Vegas:

CEE 346L: Civil Engineering Materials	Fall 2016, 2017, and 2018
CEE 370L: Engineering Mechanics of Deformable Bodies	Spring 2017 and 2018

Courses taught at Ardhi University, Dar es Salaam, Tanzania:

CEE 111: Engineering Mechanics (Statics)	Fall 2013
CEE 201: Engineering Drawing (AutoCAD)	Fall 2013
CEE 222: Traffic Engineering and Transportation Planning	Spring 2014
CEE 429: Transportation Economics	Spring 2014

Courses taught at St. Joseph University in Tanzania:

CEE 405: Transportation Engineering	Fall 2013
CEE 411: Traffic Engineering and Management	Fall 2013
CEE 418: Pavement Materials and Design	Spring 2014
CEE 420: Pavement Maintenance	Spring 2014

Courses taught at Tennessee State University:

CVEN 3200: Transportation Engineering	Fall 2011 and 2012
CVEN 4090: Traffic Engineering and Management	Spring 2012 and 2013

Mentored two graduate students:

- Thesis; Modeling Order Driver Behavior on Merging Ramps.
  University of North Florida. June 2015 - June 2016
- Thesis; HOV Lanes and Potential Effects of Traffic on Adjacent Corridors in Tennessee.
  Tennessee State University. June 2016 - December 2017

Mentored undergraduate students

Supervised 17 undergraduate students in their final year projects at Ardhi University January 2014-June 2018

Publications:

Refereed Journal Papers:

- **Kutela, B.**, & Teng, H. (2019). Modeling Transitional States of Drivers Yielding Right-Of-Way to Pedestrians at Signalized Midblock Crosswalks using a Hazard-Based Multistate Model. Transportation Research Record. https://doi.org/10.1177/0361198119841859
- **Kutela B.** and Teng H. "Prediction of Drivers and Pedestrians' Behaviors at Signalized Mid-block Danish offset Crosswalks using Bayesian Networks". *Journal of Safety Research*, Vol 69, pp 75–83, 2019.
- **Kutela B,** and Kidando E., "Towards a Better Understanding of Effectiveness of Bike share Programs: Exploring Factors Affecting Bikes Idle Duration", *American Scientific Research Journal for Engineering Technology and Sciences* (ASRJETS), vol. 29, no. 1, pp. 33-46, 2017.
- Teng, H., Puli, A., **Kutela, B**., Ni, Y. and Hu, B. "Cost and Benefit Evaluation of Graffiti Countermeasures on the Nevada Highways". *Journal of Transportation Technologies*, Vol 6, pp 360-377, 2016 doi: 10.4236/jtts.2016.65031.
- Chimba D., and **Kutela, B.** "Scanning secondary derived crashes from disabled and abandoned vehicle incidents on uninterrupted flow highways". *Journal of Safety Research (*ELSEVIER), Vol 50, pp 109–116, 2014.
- Chimba D., **Kutela, B.,** Ogletree, G., Horne, F., and Tugwell, M. "The Impact of Abandoned and Disabled Vehicles to Freeway Incident Durations". *ASCE Journal of Transportation Engineering,* Vol 140 Issue 3, 2014, 04013013.
- Chimba D., Sando T., Kwigizile V., and **Kutela B.** "Modeling School Bus Crashes Using Zero Inflated Model". *Journal of Transportation Statistics* Vol 10(1), 2014: ISSN 1094-8848.
- Chimba, D., Emaasit D., and **Kutela, B**. "Likelihood Parameterization of Bicycle Crash Injury Severities". *Journal of Transportation Technologies*, Vol. 2 No. 3, 2012, pp 213-219.
- Chimba, D., Emaasit D., and **Kutela, B**. "Integrating Origin-Destination Survey and Stochastic User Equilibrium: A Case Study for Route Relocation". *Journal of Transportation Technologies*, Vol. 2 No. 4, 2012, pp 297-304.

Under Review

- **Kutela B**. and Teng H. "Assessment of Methodological Alternatives for Modeling the Spatiotemporal Crossing Compliance of Pedestrians at Signalized Midblock Crosswalks". *ASCE Journal of Transportation Engineering*, 2019 (2nd review).
- **Kutela B.** and Teng H. "The Influence of Campus Characteristics, Temporal Factors, and Weather Events on Campus-Related Daily Bike-Share Trips", *Journal of Transportation Geography, 2018* (2nd review)
- **Kutela B**. and Teng H. "Exploring the Associated Factors for Partial-yield and Near-Miss Incidents at Signalized Midblock Crosswalks". *Journal of Transportation Safety and Security,* 2018 (1st review).
- **Kutela B.** and Teng H. "Evaluating the Influential Factors for Pushbutton Utilization at Signalized Midblock Crosswalks". *Safety Science*, 2018 (1st review).

Refereed Conference Papers:

- **Kutela B.** and Teng H. Modeling Transitional States of Drivers Yielding Right-Of-Way to Pedestrians at Signalized Midblock Crosswalks Using a Hazard-Based Multistate Model. *Presented at Transportation Research Board 2019 Annual Meeting. Paper # 19-0602*
- **Kutela B.** and Teng H. Evaluating the Influence of Regulatory Signs and Audible Devices on Pedestrians' Crossing Behaviors at Signalized Midblock Crosswalks. *The Fall Transportation Conference, 2018.*
- **Kutela B.** and Teng H. Parameterizing the Yielding Compliance of Motorists at Signalized Midblock Crosswalks Using Mixed Effects Logistic Regression. *Presented at Transportation Research Board 2018 Annual Meeting. Paper # 18-02289.*
- Teng H., **Kutela B.,** and Bingyi Hu**.,** Assessment of the Effectiveness of the Dynamic Message Signs on the Freeways in Las Vegas, Nevada. *The Fall Transportation Conference*, 2016.
- Teng H, Mulokozi E., and **Kutela B.,** Feasibility Study of Bike Sharing Program on College Campuses. *The Fall Transportation Conference*, 2015. (Received the best paper award)

- Teng H., **Kutela B.** Technical Feasibility Study of Passenger Rail Service along the West Route between Las Vegas and Los Angeles. ITE Western District Annual Meeting, Las Vegas, Nevada - July 19 to 22, 2015.
- Chimba, D., **Kutela, B.,** Ogletree, G., Horne, F., and Tugwell, M. Disabled and Abandoned Vehicle Incidents Governing Laws, Literature and Survey from Different States. Presented at Transportation Research Board 2014 Annual Meeting. Paper # 14-0023.
- Chimba, D., **Kutela, B.**, Ogletree, G., Horne, F., and Hallavant, R. The Impact of Abandoned and Disabled Vehicles to Freeway Incident Durations. Presented at Transportation Research Board 2013 Annual Meeting. Paper # 13-1065.
- Chimba, D., and **Kutela, B**., Sando T., and Kwigizile V. Paralleling the Influence of Unscheduled and Scheduled Roadwork Characteristics to Traffic Incident Durations. Presented at TRB 2013 Annual Meeting. Paper # 13-0637.

Presentations

- **Kutela B**. and Teng H., Evaluation of Pedestrian-Driver Interactions and Associated Risks at Signalized Midblock Crosswalks in Las Vegas. *Nevada. Nevada safety Summit*, 2018.
- **Kutela B.** and Teng H. Evaluating the Influence of Regulatory Signs and Audible Devices on Pedestrians' Crossing Behaviors at Signalized Midblock Crosswalks. *The Fall Transportation Conference, 2018.*

Technical reports:

- Teng H., **Kutela B.,** Mulokozi E., Hu B., Jiao Y., and Li H. "Feasibility Study of a Campus-Based Bikesharing Program at UNLV" *Mineta Transportation Institute Publications* (2017)
- Teng H., **Kutela B.,** Hu B. "Field Test of Slow-Moving Traffic Alerting System on Freeways in Las Vegas, Nevada: Assessment of the Effectiveness of the Dynamic Message Signs on the Freeways in Las Vegas, Nevada" Nevada Department of Transportation. https://rosap.ntl.bts.gov/view/dot/35044, (2015)

Research Interests:

- Traffic Safety and Operations
- Human Behavior in Transportation
- Intelligent Transportation System
- Shared and Micro-mobility
- Application of Machine Learning and Artificial Intelligence in Transportation
- Connected and Autonomous Vehicles Safety
- Railroad Network Analysis
- High speed rail planning and operation

Research and Grant Proposal Experience:

Dissertation

- Modeling Driver-Pedestrian-Infrastructure Interactions at Signalized Midblock Crosswalks
    - Developed multistate models for modeling yielding compliance of drivers at signalized midblock crosswalks by considering the transitional states and the corresponding transitional durations.
    - Presented the methodological framework to assess the modeling methodologies for spatiotemporal crossing compliance of pedestrians at signalized midblock crosswalk. This proper approach was deemed since both spatial and temporal crossing compliance occur jointly
    - Proposed and presented the methodology to model near-miss and multiple threats incidents which are rarely observed at signalized midblock crosswalks but are very risky.
    - Demonstrated the application of supervised machine learning techniques in evaluating the combined influence of crosswalk features on prediction of drivers' spatial yielding compliance.
    - Evaluated the influential factors for pushbutton utilization at signalized midblock crosswalks
    - Evaluated the influence of regulatory signs and audible devices on pedestrians' crossing behaviors at signalized midblock crosswalks

Grant proposal preparation.

- As a Graduate Assistant I took place in preparing one research proposal titled "Prioritization of Wildlife-Vehicle Conflicts in Nevada" under my supervisor Dr. Hualiang Teng.

Research projects

- High Speed Rail Access Charge for the XpressWest of Nevada, under Jin Ouk Choi (PI), Mohamed Kaseko (Co-PI), and Hualiang (Harry) Teng (Co-PI).
- Developing and testing an LED system to improve pedestrian safety in Nevada, funded by Nevada Department of Transportation (NDOT), under Dr. Hualiang Teng (PI), at University of Nevada Las Vegas.
- Feasibility Study of Bike Sharing Program on College Campuses, funded by Nevada Department of Transportation (NDOT), under Dr. Hualiang Teng (PI), at University of Nevada Las Vegas.
- Technical Feasibility Study of Passenger Rail Service along the East Route between Las Vegas and Los Angeles, self-funded with software donation from Berkeley Simulation Software, LLC. and Las Vegas Railway Express, Inc., under Dr. Hualiang Teng (PI), at University of Nevada Las Vegas.
- Field Test of Slow-Moving Traffic Alerting System on Freeways in Las Vegas, Nevada: Assessment of the Effectiveness of the Dynamic Message Signs on the Freeways in Las Vegas, Nevada, funded by the Nevada Department of Transportation (NDOT), under Dr. Hualiang Teng (PI), at University of Nevada Las Vegas.
- Statistical Evaluation of Abandoned and Disabled Vehicles on Tennessee Highways, sponsored by the Tennessee Department of Transportation (TDOT), under Dr. Deo Chimba (PI), at Tennessee State University

My roles in the projects: Laboratory and site testing of the system, attend site meetings, survey questionnaire preparation and online distribution, literature review, data collection and analysis, simulation of the railroad network, and report preparation.

Engineering and Data Analytics Skills:

- **Engineering related software:** AutoCAD Civil 3D, HCS, Microstation and TrafficWare Synchro, ArcGIS and QGIS Rail Traffic Controller
- **Programming Languages & Big Data:** Python, R, Hadoop, SparkR, TensorFlow, SQL
- **Machine Learning:** Logistics Regression, Naive Bayes, Decision Tree, Random Forest, KNN, Linear Regression, SVM, Regression Tree, K – means, Bayesian Networks, Artificial Neural Networks.

Organizational skills:

Co-organized eight Seminars/Symposiums

- Seminar on Railroad Dynamics, University of Nevada, Las Vegas October 2018
- Railroad Infrastructure Diagnosis and Prognosis Symposium,
  University of Nevada, Las Vegas October 2018
- High-Speed Rail Seminar, University of Nevada, Las Vegas December 2017
- Lessons for the United States from High-Speed Rail's Urban Impact in China.
  University of Nevada, Las Vegas September 2017
- Fundamentals and Selected Technical Issues for High Speed and Heavy Axle
  Railroad Engineering. University of Nevada, Las Vegas June 2017
- Deterioration modeling of rail infrastructure: The factor moment Approach.
  University of Nevada, Las Vegas March 2017
- AREMA 3-day seminar: Introduction to Practical Railway Engineering, Las Vegas May 2016
- Public Transportation Systems, University of Nevada, Las Vegas December 2015

Awards and Scholarships:

- **Travel Grant**, 98th Annual Meeting of the Transportation Research Board,
  Washington DC January 2019
- **Travel Grant** |Transportation Technology Center Inc University Day, Pueblo, Colorado July 2018
- **Travel Grant**, 97th Annual Meeting of the Transportation Research Board,

Washington DC January 2018
- **Graduate Teaching Assistantship**, University of Nevada Las Vegas August 2016 – to present
- **Graduate Research Assistantship**, University of Nevada Las Vegas August 2014 –May 2016
- **UNLV Access Grant**, University of Nevada Las Vegas August 2015 – August 2018
- **Travel Grant**, International Seminar on High Speed Technology on Railway System, San Jose, California June 2015
- **Best Student Paper and Oral Presentation**, Fall Transportation Conference, Las Vegas, NV October 2015
- **Graduate Research Assistantship**, Tennessee State University, Full Master's Scholarship August 2011 – May 2013
- **Best Graduating Engineer in Structural Engineering,** Engineers Registration Board (Tanzania) September 2011
- **Best Final Year Student in Civil and Structural Engineering,** University of Dar es Salaam October 2010
- **Best Student in Structural Analysis,** University of Dar es Salaam October 2010
- **Best Student in Structural Design,** University of Dar es Salaam October 2010
- **Best Third Year Student in Civil and Structural Engineering,** University of Dar Es Salaam October 2009
- **Best Second Year Student in Civil and Structural Engineering,** University of Dar Es Salaam October 2008

Professional Involvement:

Journal Reviewer:

- Journal of Transport and Land Use May 2015 to present

Professional Membership and Leadership involvement:

- Represented our UTC center in the CUTC Winter meeting in Washington DC January 2019
- Co-Organizer, Data Science Las Vegas July 2015 to present
- Student Member, American Society of Civil Engineers (ASCE) March 2016 to present
- President of AREMA student chapter at UNLV August 2016 – May 2018
- Student Member, American Railway Engineering and Maintenance-of-Way Association August 2015 to present
- Student Member, Institute of Transportation Engineers (ITE) January 2012 to present
- Member, Institution of Engineers Tanzania July 2010 to present

Industry Experience:

Civil Engineering Intern: Dar es Salaam, Tanzania
Teknicon Ltd. May 2010- July 2010

- Structural design and detailing of residential and commercial buildings

Civil Engineering Intern: Dar es Salaam, Tanzania
NIMETA Consulting Engineers and Planners March 2010- April 2010

- Structural design and detailing of residential and commercial buildings
- Preparation of technical and financial proposals

Civil Engineering Intern: Mpanda, Tanzania
Works Department June 2009- Sept 2009

- Structural design and detailing of bridges
- Reviewing technical drawing for building permits
- Site visits and other management work