EFFECTIVENESS OF HIGH-VISIBILITY CROSSWALKS ON PEDESTRIAN SAFETY SURROGATES: AN EXPLORATORY EMPIRICAL ANALYSIS USING THE SHRP 2 NATURALISTIC DRIVING DATA

By

Courtney Bentley

Graduate Research Assistant

Department of Civil, Structural, and Environmental Engineering

Engineering Statistics and Econometrics Research Laboratory

University at Buffalo, The State University of New York

204B Ketter Hall, Buffalo, NY 14221, (716) 645-2114, cebentle@buffalo.edu

M.Sc. Thesis

State University of New York at Buffalo

Buffalo, New York

December 3, 2015

ACKNOWLEDGEMENTS

I would like to express my gratitude to my advisor, Professor Panagiotis Ch. Anastasopoulos, for his support and guidance in completing this thesis. My thanks also go to my committee member, Professor Adel Sadek, for his help and input. I would like to thank the NYSDOT and CUBRC for their partnership on the SHRP2 Implementation Assistance Program (IAP) Phase 1 project that this began as part of, "Do High Visibility Crosswalks (HVCs) Improve Pedestrian Safety?" I also want to thank the FHWA for sponsoring the SHRP2 IAP. I am grateful to my colleagues, Tawfiq Sarwar and Ugur Eker, for their countless hours spent helping me to watch videos and record data. I want to also thank my professors and colleagues at the University at Buffalo for their support throughout my college career, as well as my coworkers at Bergmann Associates for their patience and understanding as I worked part-time. My appreciation goes to all of my family and friends, especially my parents, who have supported me through my graduate studies and completing this thesis.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

This thesis evaluates the effectiveness of high-visibility crosswalks (HVC) to improve pedestrian safety at uncontrolled locations using the SHRP2 naturalistic driving study (NDS) data. This is accomplished by analyzing the driving behavior of SHRP2 participants at three uncontrolled locations in the Erie County, New York test site. At two intersections, traversal data were available both before and after HVC installation allowing for a before/after analysis. At the third location, only post HVC installation data were available. Because no pedestrian – motor vehicle crashes were observed, crash surrogates (i.e., speed, acceleration, and gas pedal position) were used to evaluate driving behavior. Random effects linear regression models with fixed and random parameters were estimated for the change in the surrogate measures between predetermined benchmark points and the crosswalks, while controlling for a variety of other factors. The results show that presence of both the HVC and pedestrian crossing sign decreased the change in speed and acceleration between the benchmark and crosswalk points. In addition, there was a greater deceleration between the benchmark and crosswalk points after HVC installation. This exploratory work shows that HVCs have the potential to improve pedestrian safety and modify driving behavior, and that NDS data are useful for analyzing their effectiveness.

CHAPTER 1. INTRODUCTION

1.1. Background and Motivation

In the United States, traffic crashes injured an estimated 66,000 pedestrians and killed 4,735 in the year 2013. These pedestrian deaths accounted for 14 percent of all fatalities in traffic crashes. In New York State alone, pedestrian fatalities accounted for 28 percent of the total fatalities on the state's roadways (NHTSA, 2013; NYS, 2014). Making roadways safer for pedestrians is an important national and statewide goal. Many strategies to accomplish this goal have been studied, such as: passive markings and signage (e.g. high-visibility crosswalk markings), traffic calming measures (e.g., roadway narrowing, horizontal shifts, and vertical deflections), and active control devices (e.g., automated pedestrian detection, smart lighting, and high intensity activated crosswalks). It is important to carefully consider which countermeasure will be the most effective at a given location, including whether marked crosswalks should be provided at all.

Safety analysis of a location is often retroactive, that is, after accidents have occurred and pedestrians have been injured or killed. Safety countermeasures are either implemented as a reaction to accidents, or proactively without adequate studies specific to the location or countermeasure type. It is diffult to accurately analyze driver behavior and reactions to various countermeasures. Past research has used various methods from driving simulators to field observations of speed. While driving simulators offer resarchers an opportunity to observe drivers under controlled conditions, it does not fully represent how drivers would behave in their own vehicle on the roadway. In-vehicle field observations can address some of these limitations, but drivers may behave differently with an observer in the car. Speed and yielding field observations are useful to measure overall effect of safety countermeasures, as researchers can study how many drivers are reacting to certain conditions. However, specific behavior characteristics can not be observed and anlyzed. Naturalistic driving study (NDS) data provide a unique opportunity to analyze driver behavior, as details about every day trips are recorded for an extended period of time without experimental control. Drivers' reactions to crosswalk striping and other safety countermeasures can be accurately measured and anlyzed.

1.2. Objectives

The present study focuses on the relatively low cost, widely used pedestrian safety strategy of high-visibility crosswalk (HVC) markings. The overall goal is to evaluate the effectiveness of HVCs to improve pedestrian safety at uncontrolled locations. This is accomplished using the 2nd Strategic Highway Research Program (SHRP2) naturalistic driving study (NDS) data collected at the Erie County, New York test site. Driving behavior of SHRP2 participants is analyzed at three uncontrolled locations with HVC markings. For two of the locations, traversal data was acquired both before and after HVC installation, which allowed for before-after analysis. At the third location only post-HVC installation data was available. Statistical models are developed to evaluate driver reactions prior to crossing the HVC while controlling for a variety of factors, including intersection and roadway geometric characteristics; traffic characteristics; time and date of trips; lighting, pavement, and weather conditions; driver characteristics; and vehicle characteristics.

No pedestrian – motor vehicle crashes were observed in the Erie County SHRP2 test site; therefore, pedestrian safety surrogates are analyzed in the before-after analysis and statistical modelling. The surrogate measures (i.e., speed, acceleration, and gas pedal position) represent driver behavior and have an impact on pedestrian safety at crosswalks. The parameters used are the values of the three measures at a benchmark point prior to the crosswalk, at the HVC location, and the difference between the two points. The desirable effects to improve pedestrian safety are shown in Table 1.1. In general, slower speeds, less acceleration (or more deceleration), and lower gas pedal position (less pressure) can increase pedestrian safety.

Table 1.1. Desirable effects of HVCs on surrogate measures to improve pedestrian safety.

Parameter	**Desirable Effect for Pedestrian Safety**
Speed at Benchmark *(km/h)*	Slower speed
Speed at HVC *(km/h)*	Slower speed
Speed Difference Between Benchmark and HVC *(km/h)*	Decrease (more slowing between benchmark and HVC)
Acceleration at Benchmark *(g)*	Decrease (less acceleration or more deceleration)
Acceleration at HVC *(g)*	Lower (less acceleration or more deceleration)
Acceleration Difference Between Benchmark and HVC *(g)*	Decrease (greater deceleration between benchmark and HVC)
Gas Pedal Position at Benchmark	Lower (less pressure on gas pedal)
Gas Pedal Position at HVC	Lower (less pressure on gas pedal)
Gas Pedal Position Difference Between Benchmark and HVC	Decrease (let up on gas pedal between benchmark and HVC)

CHAPTER 2. LITERATURE REVIEW

2.1. Introduction

Determining the best methods to increase pedestrian safety has been the topic of many engineering evaluations and research projects. Building infrastructure for pedestrian facilities is necessary to provide mobility and allow people to cross the road safely, especially in urban and built-up areas where pedestrian volumes are high. Pedestrian crossings at unsignalized intersections and mid-block locations are a top safety concern. There are various factors that affect how effective a countermeasure will be at a given location. The effectiveness of high-visibility crosswalk (HVC) markings, particularly at uncontrolled locations, has been a topic of debate in the traffic safety community. Past research evaluating crosswalks has produced varying results. There are many approaches to evaluating safety, from analyzing accident rates to statistical analysis of crash or speed data.

2.2. Pedestrian Crash Analysis

In the late 1960s, Herms analyzed crashes at marked and unmarked crosswalks at unsignalized intersections in San Diego, California. He found that there were significantly more pedestrian crashes in marked crosswalks as in their unmarked counterparts (Herms, 1972). The study compared accident counts and did not take into account the effects of other factors that could have led to increased pedestrian crashes. The results of the study lead to the removal of many unsignalized pedestrian crossings in urban areas, which has been criticized for limiting pedestrian mobility.

Another study by Jones and Tomcheck (2000) analyzed vehicle-pedestrian collisions at uncontrolled intersections in Los Angeles, California where marked crosswalks were not reinstalled after roadway resurfacing. The before-and-after study of crash history showed that the number of pedestrian-vehicle collisions decreased where markings were removed, and crashes at adjacent intersections where markings were reinstalled only increased slightly showing that the crashes at removed crosswalks were not being transferred to adjacent marked crosswalks.

Zegeer et al. (2005) performed a more comprehensive analysis of pedestrian accidents at uncontrolled crosswalks by accounting for other factors such as speed, traffic volume, and street width. The study analyzed pedestrian crashes at a set of uncontrolled marked crosswalks and unmarked comparison sites. The findings show that as number of lanes, traffic volume, and speed limit increase, crosswalk markings alone are related to higher crash frequency compared to locations with no markings. However, raised medians on multilane roads resulted in a lower crash rate. One explanation given for increased crashes at crosswalks is that installing marked crosswalks leads pedestrians to choose to cross at the uncontrolled location rather than using the closed signalized crosswalk.

Therefore, marked crosswalks increased the number of at-risk pedestrians. The study recommended that other treatments should be installed in addition to crosswalk markings to provide a safer street crossing than a crosswalk alone. The analysis by Zegeer et al. was based on a small amount of data with an average of one pedestrian crash per crosswalk site every 43.7 years (Zegeer et al., 2005, pg. 16). Over the average of 5 years of data across the 2,000 sites, there were only 229 pedestrian crashes. Problems of past studies were addressed by including other roadway and traffic factors, but the few observations of crashes indicates that the factors that affect crashes may not be accurately represented.

Studies have been completed to determine the factors that affect pedestrian injury severity at unsignalized crosswalks. Haleem, Alluri, and Gan (2015) reported that factors including darkness, increased speed limit, and pedestrians walking along the roadway resulted in higher pedestrian severity risk. Similarly, Olszewski et al. (2015) found that divided roads, two-way roadways, mid-block locations, darkness, and higher speed limits increase pedestrian fatality risk. These results suggest reducing crash severity by improving roadway lighting near crosswalks, avoiding installing crosswalks on high speed roads, and signalizing crosswalks on divided roads.

2.3. Driving Behavior

In addition to analyzing accident history, research has been done evaluating drivers' reactions and their speed at crosswalks. It has been found that crosswalk markings generally result in decreased vehicle speeds (Knoblauch & Raymond, 2000; Knoblauch, Nitzburg, & Seifert, 2001). Nitzburg and Knoblauch (2001) evaluated the effectiveness of HVC markings combined with pedestrian crossing signs and reported an increase in driver yielding behavior. However, these studies looked at average driver behavior before and after HVC installation, which cannot capture actual behavior of individual drivers. They also reported an increase in crosswalk usage by pedestrians after the installation, which supports the observation that pedestrians prefer to use crossing locations that have marked crosswalks.

To study individual drivers, driving simulator studies have also been used to evaluate driver behavior and pedestrian safety. Gómez et al. (2013) analyzed advance yield markings for a marked midblock crosswalk. The markings are a row of triangles painted on the raodway before the crosswalk, with signs telling drivers to yield there for pedestrians. The study measured perfomance with crashes and glancing behavior. With the advance yield markings, fewer crashes occurred and drivers looked more frequenty and sooner for pedestrians at the crosswalk. The project was validated with a field evaluation of advance yield markings by Samuel et al. (2013). The study involved staged pedestrians attempting to cross and an in-vehicle field study on an open road course. In addition to validating the driving simulator results, it was determined that vacating parking spots adjacent to the crosswalk (at least between the yield markings and crosswalk) improved yielding behavior by clearing the sight line.

The previous research was used by Fitzpatrick et al. (2006) to develop guidelines for selecting the appropriate pedestrian crossing treatments at unsignalized intersections and midblock locations. Motorist compliance (i.e., yielding or stopping) was considered the measure of effectiveness, as it has been shown to be related to pedestrian safety. The guidelines include the roadway and traffic characteristics that should be considered in determining which specific treatment(s) will provide a safe crossing when crosswalks are being installed.

High-visibility crosswalk (HVC) marking styles were reviewed by McGrane and Mitman (2013) to determine what types of crosswalk markings were the most effective. They concluded that high-visibility markings are more easily detected by motorists than other styles, so that drivers become aware of the potential for pedestrians to be present and can yield sooner. Whenever it is determined that marked crosswalks should be provided at an uncontrolled location, HVC markings should be installed. While they can improve overall visibility of a crosswalk, HVCs are the most effective when combined with other enhancements such as warning beacons, signage, or geometric improvements.

2.4. Methodological Approach

Many different statistical modeling approaches have been used to analyze roadway and traffic safety data. One common approach is to study accident counts, as discussed above. Count data are more accurately analyzed if they are normalized into accident rates to compare across locations and times. Accident counts are commonly normalized by traffic volumes. Accident rates are continuous data that can only take positive values, and are left-censored at zero because there are many roadway segments with no observed accidents over a given period of time. For these reasons, Tobit regression analysis is an applicable model form to use to analyze accident rates.

Tobit analysis has been applied in past research, such as by Anastasopoulos et al (2008). The application of Tobit models was explored by applying it to accident rates on interstate segments in Indiana. By using a modeling approach that could account for left-censored data, the complete available data could be used by including segments with and without observed accidents. The results showed that a variety of pavement, geometric, and traffic factors had a significant effect on vehicle accidents, and that the Tobit model fit the data well. Another study by Anastasopoulos et al (2012e) incorporated random parameters to account for unobserved heterogeneity. The same data from Indiana was used, and the results were compared. It was shown that the random parameters Tobit model outperforms its fixed parameters counterpart.

Tobit regression analysis was also applied to highways in Washington State by Anastasopoulos et al (2012f), and took a new approach by considering accident rates by injury-severity (i.e., no-injury, possible injury, and injury). The multivariate Tobit model that was applied was found to outperform its univariate counterpart. The effects of exploratory parameters were not the same across the injury-severities, which shows that a multivariate approach has the potential to provide a fuller understanding of the factors of accident rates and injury-severities.

The factors that affect accident rates have also been analyzed by taking different approaches. Anwaar et al (2012) used aggregate data for 178 countries to study the relationship between traffic safety, health service levels, and motorization levels. The data that were used were the first comprehensive set of data made available by the World Health Organization and the International Road Federation, and the factors were normalized by population to compare across countries of different sizes. Two modeling specifications were tested to estimate the national roadway fatality rate: a set of regression models, and a system of seemingly unrelated regression equations (SURE) models. It was determined that the SURE model was statistically superior to the separately-estimated regression models. The results show that a higher vehicle fatality rate in a nation is associated with factors such as low-to-median income levels, low road network density, and low enforcement of seatbelt laws. The data has limitations related to different measurement techniques between countries, especially developing countries; however, the results offer preliminary insights on identifying nationwide patterns to address in order to improve traffic safety.

Factors that affect accidents vary across locations, resulting in unobserved heterogeneity. Instead of assuming that parameters are fixed across observations, random parameter statistical modeling allows some parameters to vary. Anastasopoulos and Mannering (2009) applied random parameters to accident frequency count models and demonstrated that random parameters regression has considerable potential for analyzing accident data. Ignoring the possibility of unobserved heterogeneity can result in significantly different results. Random parameters were also applied by Anastasopoulos and Mannering (2011) to crash-severity logit models to estimate the severity of a crash given that it has happened. When the statistical fit of the models was compared to the traditional fixed parameter models, it was found that the random parameter models provide a statistically better fit.

Random parameter modeling has been used to improve analysis accuracy with other modeling approaches and in areas other than traffic safety, such as pavement condition, travel times, and project contract types. Anastasopoulos et al (2011b) explored state-level pavement performance using logit models. Aggregate state-level data on pavement performance, money spent on preservation, surface geology, and climate were used to estimate logit models for pavement condition by roadway functional class. Random parameters were incorporated to account for random variations across geographic locations. The exploratory study demonstrated the potential for using random parameter models to analyze pavement performance. The potential was again demonstrated with the seemingly unrelated regression equations (SURE) approach. Anastasopoulos et al (2012d) used random parameters SURE models to determine the performance of rehabilitated pavements. The SURE approach was used to account for the interrelation among the various pavement performance measures used.

Random parameters have also been used to account for unobserved heterogeneity in hazard-based duration models, such as by Anastasopoulos and Mannering (2015). Using a forecast with SURE models and historical pavement condition thresholds, the service life of pavement overlays and replacement were determined. Random parameter duration models were then estimated to identify the influential factors of overlay and replacement performance. When compared to corresponding fixed parameters models, the random parameter models were found to have a statistically superior fit of the data at a 99.9% confidence level. Random parameters duration models were also used to assess the likelihood and duration of highway project delays by Anastasopoulos et al (2012c). A random parameters binary logit model was used to evaluate the likelihood of time delay based on project characteristics, then the duration models were used to study the factors that contribute to the duration of the project delay.

Random parameters with a hazard-based approach has also been used to analyze urban travel times (Anastasopoulos et al, 2012a). The traditional approach to travel time analysis uses a complex modeling system based on activity and trip generation methods. A more simplistic approach was used to focus on the travel time data alone, but simplifying such a complex decision-making process introduces unobserved heterogeneity with effects that vary across the population. This was accounted for by using random parameters for many of the characteristics that were found to be significant, including socio-demographics, trip characteristics, travel mode, and time of day of the trip. Analysis of the results showed that the random parameters model had a better statistical fit than its fixed parameter counterpart.

Incorporating random parameters with various modeling methods was found to improve statistical fit in other studies analyzing the performance of highway project contract types including public-private partnerships and performance-based contracts (Anastasopoulos et al; 2009, 2010a, 2010b, 2011a, 2014). Based on the results of these and the research that was discussed in more detail above, it can be determined that random parameters offer a good opportunity to account for possible estimation issues. Rather than the traditional approach of assuming that parameters have the same effects across the population, random parameters can account for the unobserved factors that can cause the effects to vary.

2.5. Summary

Previous research has shown that marked crosswalks at uncontrolled locations can potentially increase pedestrian accidents in some locations, but improve safety in other locations. The effectiveness of crosswalk markings implemented alone depends on the specific roadway and traffic characteristics of the location. Some studies did not fully and accurately incorporate these other factors, and the results led to the removal of crosswalks. However, this means decreasing pedestrian mobility and the lack of crossing facilities could lead to more severe crashes if pedestrians choose to cross at that location anyway because other crossings are much farther away and more inconvenient.

When implemented with other safety countermeasures, crosswalk markings can improve pedestrian safety by increasing driver awareness of potential pedestrians and decreasing vehicle speeds. These countermeasures could include geometric improvements, high visibility markings, and warning signs. High-visibility crosswalk (HVC) markings have been found to be more easily detected by driver than other styles, and is the best option when the decision is made to provide a crossing at uncontrolled locations. However, HVCs perform better at improving safety when combined with other enhancements such as pedestrian crossing signs. This combination has been found to increase driver yielding before the crosswalk. While these results could help policy makers analyze when, where, and what type of crosswalk can improve pedestrian safety, the past studies had limitations. The data and methods that were used could not fully capture driving behavior that occurs in the field to analyze HVC effectiveness.

The SHRP2 naturalistic driving study (NDS) provides a great opportunity to address the limitations of past research. The data is a rich source of information that captures individual driving behavior rather than aggregated average values. NDS data gives a complete representation of actual driver behavior, unlike driving simulators or field studies with an observer in the vehicle. Drivers were observed over an extended length of time, which allows for more analysis of their behavior and reactions than a driving simulator or field test. With the SHRP2 data, other factors can be accounted for including gender and age differences, environmental conditions, and trip frequency by the same driver. Statistical modeling of measures of driving behavior accounts for other factors that could impact pedestrian safety. This study will address past issues by analyzing actual driving behavior before and after the installation of HVCs at uncontrolled locations.

CHAPTER 3. DATA AND METHODOLOGY

3.1. Overview of Approach

The Naturalistic Driving Study (NDS) for the second Strategic Highway Research Program (SHRP2) was conducted with main purpose of understanding the role of driver performance and behavior in roadway safety. Unlike the data used in the past, the NDS data offers detailed information on the everyday driving behavior of a large number of participants. This exploratory analysis uses a sample of the data from one of the six NDS sites, to evaluate the effectiveness of high-visibility crosswalks (HVC). Three HVC locations were selected for the study, and a representative random sample of trips through the locations were analyzed. The data used includes forward-facing videos and time series data for each trip, as well as basic driver and vehicle characteristics.

No pedestrian – motor vehicle crashes were observed; therefore crash surrogate measures (i.e., speed, acceleration, and gas pedal position) were used to evaluate driving behavior. The surrogate measures were selected based on factors that have been found to impact accidents in past research, and are representative measures of driving behavior. For the sites with available data before and after the HVCs were installed, hypothesis tests were performed to determine if there was a significant change in the surrogate measures with the implementation of HVCs. While hypothesis tests indicate if the HVC improve pedestrian safety, they do not take into account other factors that affect driving behavior. Linear regression models were estimated for the change in the surrogate measures between a benchmark point and the crosswalk while controlling for a variety of other factors.

3.2. Study Location

Trip data were collected and processed from three representative high-visibility crosswalk (HVC) locations in the Erie County SHRP2 study site. The data were collected over the three-year period from 2011 to 2013. Sites were chosen based on the availability of sufficient traversal data through the locations both before and after the HVC was installed. However, one of the locations only provided trips after the HVC was installed, but was selected because it had a relatively high pedestrian traffic count compared to the other locations. This allowed for the statistical models to account for the effect of pedestrian presence on driver behavior.

The HVCs at all three locations consisted of ladder-type crosswalk markings, as well as pedestrian crossing signs. There were warning signs installed in advance of the crosswalk, as well as at the crossing location. Table 3.1 identifies the crosswalks and summarizes information about the three locations. Two of the locations (Eagle Street with Oak Street and Elm Street, in Buffalo) are 3-lane one-way streets that run parallel to each other but in opposite directions, and both are stop-controlled on the minor approach (Eagle

Street). These two Buffalo locations had HVC markings and pedestrian crossing signage installed during the study period. The third crosswalk is at a midblock location on Main Street in the Village of Hamburg. It is a 2-way two-lane street with on-street parking on both sides. Aerial views of the three crosswalks are shown in Figure 3.1 and Figure 3.2.

Table 3.1. Summary of selected HVC locations.

HVC Number	Name	Date Installed	No. of drivers/trips	Lanes/Direction
5	Elm/Eagle - Buffalo	8/30/12	48/474	3 lanes one direction (N)
6	Oak/Eagle - Buffalo	9/12/12	9/328	3 lanes one direction (S)
18	Main St - Hamburg	Unknown – before study period	19/276	1 lane each direction (E/W)

Figure 3.1. Aerial view of HVC locations 5 (Elm/Eagle) and 6 (Oak/Eagle) in Buffalo. (Source: Google Earth)

Figure 3.2. Aerial view of HVC location 18 (Main Street) in Hamburg. (Source: Google Earth)

3.3. Data Source

The data were collected through the SHRP2 naturalistic driving study (NDS) during the 3-year time period from 2011-2013. The study involved over 3,100 volunteer drivers who participated for a 1- or 2- year period at sites in six states. Each vehicle was equipped with an onboard data acquisition system (DAS) that included four video cameras, radar, accelerometers, vehicle network information, and GPS (FHWA, 2015). A schematic of the DAS is shown in Figure 3.3. The video cameras recorded the forward roadway view, driver's face view, downward view recording the driver's interaction with the dashboard, and a rear and right-side view. The cameras were located in the head unit mounted near the rear view mirror, shown in Figure 3.4 along with an example of the four fields of view.

In addition to the video cameras, the DAS was connected to the vehicle network and continuously recorded information while the vehicle was on. The vehicle network data includes the accelerometer, brake pedal activation, steering wheel angle, speed, seat belt information, and many other variables. The system also used radar, GPS, and accelerometers to determine the location, speed, lateral movement, and the location of surrounding vehicles (Campbell, 2012). Due to the varying vehicle models and years, the program categorized vehicles according to the amount of vehicle control data, with 'prime' and 'sub-prime' vehicles providing the maximum amount of data from the DAS. Only these two classifications were used for this project to minimize the amount of missing time series data.

In addition to the DAS, the NDS included information about each participant such as socio-demographic factors and a variety of driver assessment tests that were performed while the vehicle was being equipped. The assessment evaluated through filling in forms, computer-based tests, and physical tests provided information on visual perception, personality factors, general medical condition, and driving knowledge (SHRP2). Along with the NDS, a roadway information database (RID) contains detailed roadway data for the study sites with roadway geometry, pavement condition, traffic characteristics, accident data, and weather history.

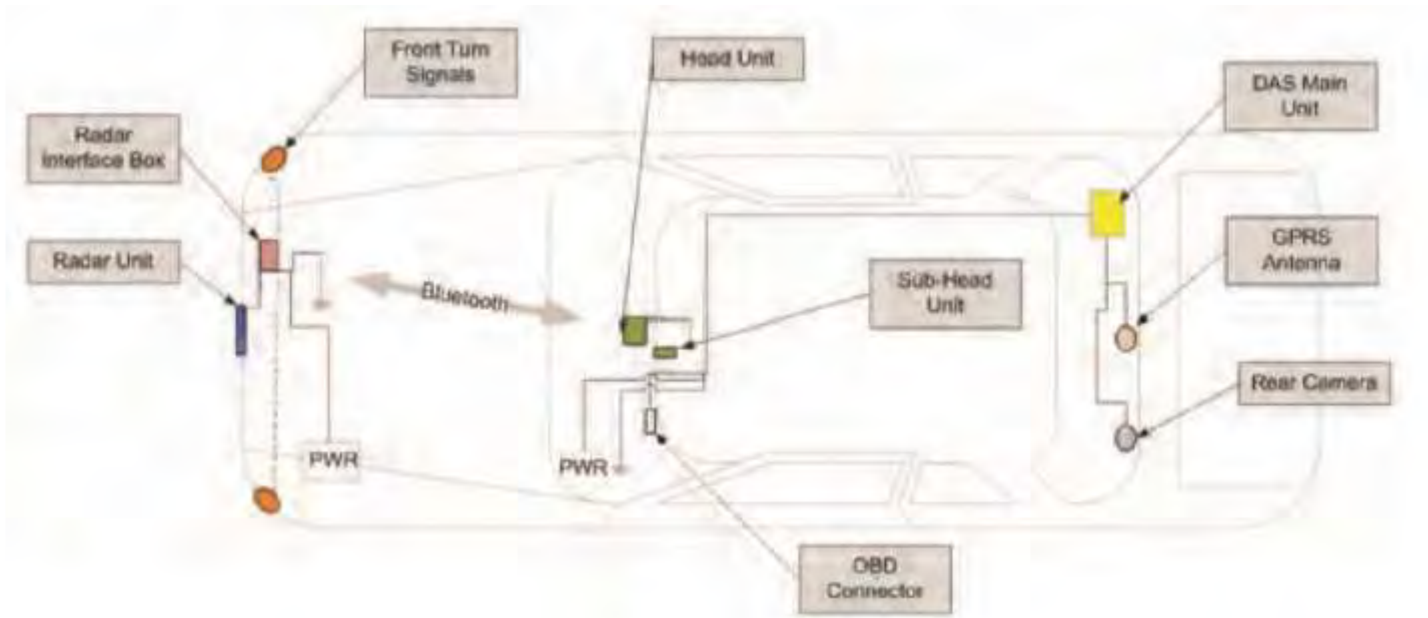

Figure 3.3. Data acquisition system schematic view (source: Campbell, 2012, page 32).

Figure 3.4. Head unit placement and camera views (source: FHWA, 2015, slide 15).

3.4. Data Selection

The data obtained through the SHRP2 NDS for this project includes information collected from the front-facing videos and time series data for each trip, in addition to basic vehicle and driver information for each participant (i.e., gender, age, and vehicle make and model). The dataset includes 1,078 trips made by a representative random sample of 62 participants across the three selected HVCs. The trips were selected to proportionally represent all gender and age groups, and both frequent and infrequent traversals of the studied locations. The sample also includes trips before and after the HVC installation for the Buffalo HVCs that were installed during the study period, as shown in Table 3.2. The total number of drivers is greater than the 62 participants because some drivers travelled through multiple locations. The sub-sample of 802 trips by 57 drivers at the Buffalo locations of 5 and 6 was used for a before-after analysis.

Although there was no data prior to the HVC installation for location 18 in Hamburg, there were more observations of pedestrians as shown in Figure 3.5Figure 3.6, and Figure 3.7. Ten percent of all trips through location 18 had pedestrians crossing the HVC, compared to only 3 to 5 percent of before and after trips through locations 5 and 6 in Buffalo. This offered a better opportunity to evaluate how drivers reacted to HVCs differently when there were pedestrians crossing or attempting to cross versus when there were no pedestrians in the vicinity.

Table 3.2. Distribution of trips before and after HVC installation

Location	**Trips Through Location**	**No. Of Drivers**	**% Trips Prior to HVC Install**	**% Trips After HVC Install**
5	474	48	42%	58%
6	328	9	58%	42%
18	276	19	0%	100%
Total	**1,078**	**76**		

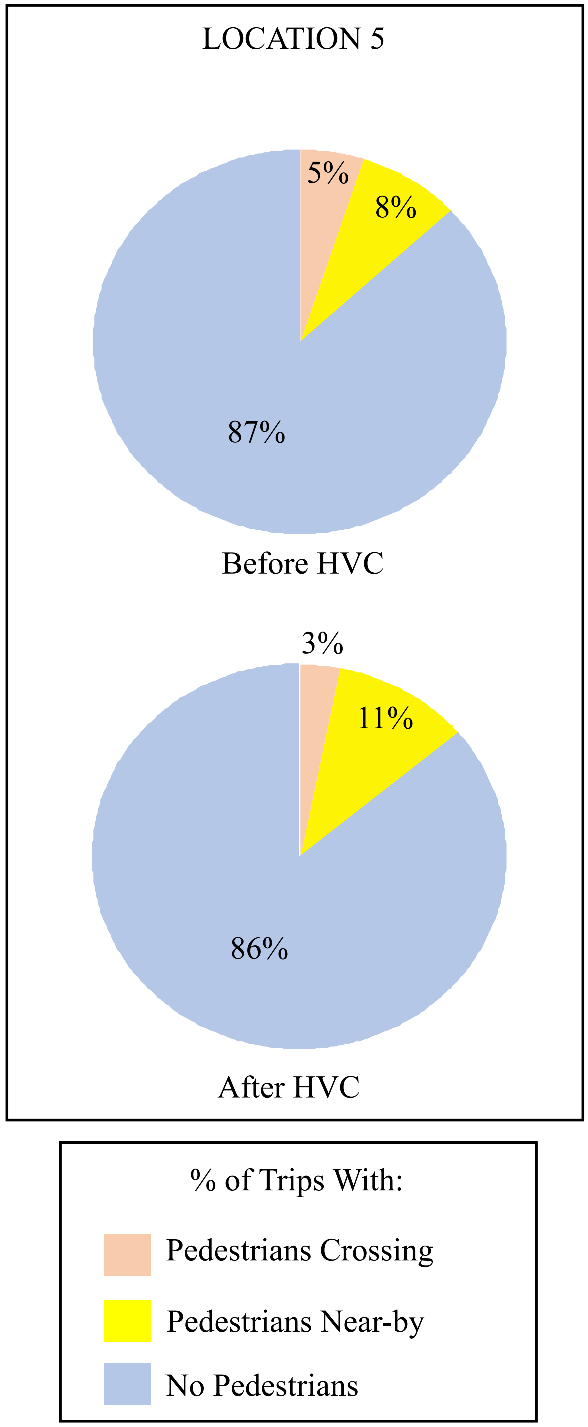

Figure 3.5. Percent of trips with pedestrians present before and after HVC installation: Location 5, Buffalo.

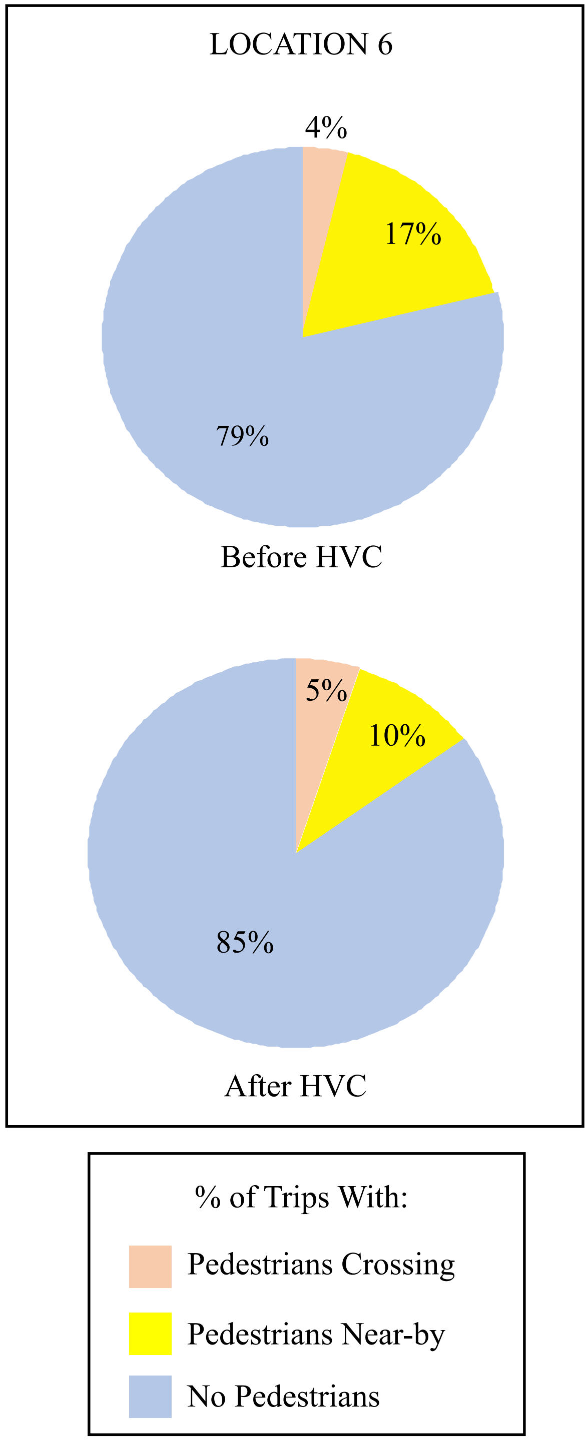

Figure 3.6. Percent of trips with pedestrians present before and after HVC installation: Location 6, Buffalo

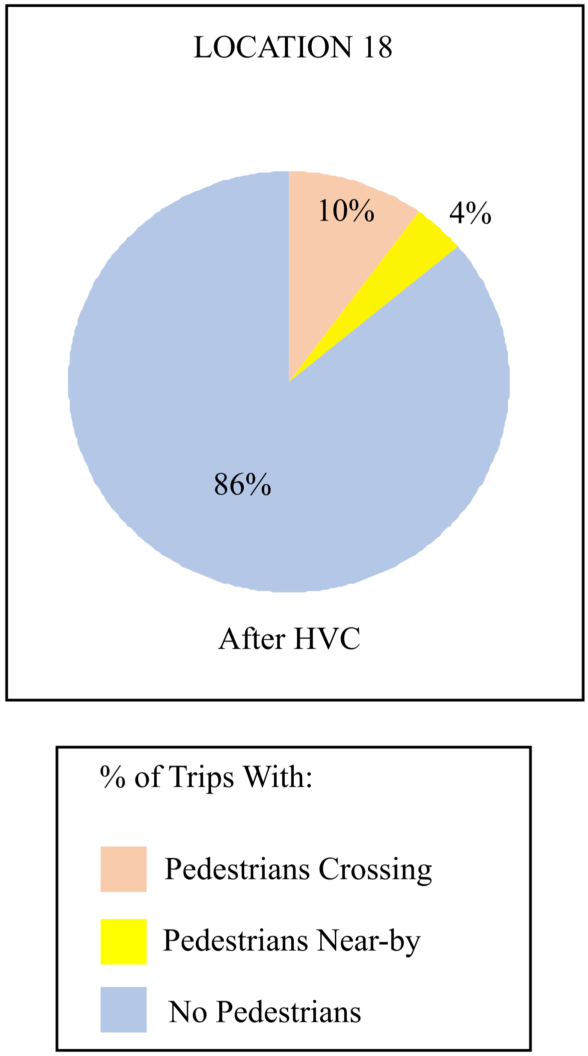

Figure 3.7. Percent of trips with pedestrians present: Location 18, Hamburg.

3.5. Data Processing

The different datasets (videos, time series data, driver, and vehicle data) were linked by the event IDs for each trip that were associated with the date, time, location, and the participant ID, which was used to identify the driver and vehicle characteristics. The time series data linked to each event ID included information recorded from the vehicle throughout the trip (e.g., acceleration, speed, brake pedal, and gas pedal position). Each trip and event ID had a corresponding forward-facing video which was cut to the time before, during, and after the HVC location traversal.

The first aspect to the analysis is converting the videos into manageable data, which involved the determination of a benchmark point for each location and direction. The benchmark points were selected to represent the approximate location where drivers are able to see and react to the HVC. They were also selected based on easily identifiable locations in the videos both before and after the HVC was installed (i.e., landmarks such as buildings and light poles were used). The videos were observed and the time that the vehicle crossed the benchmark and HVC locations were recorded. Additional information was also recorded, such as pedestrian presence, vehicle's lane position, preceding and parked vehicles' presence and the level the obstructed visibility of the HVC, windshield condition and wiper usage, weather conditions, pavement surface conditions, and lighting conditions. Using the time stamps on the video, the time series data were matched with the rest of the trip data. Since the on-board vehicle equipment records information at intervals, the exact values at the benchmark and HVC locations were interpolated.

The selection of the benchmark points for all locations was based on the stopping sight distance (i.e., the reaction distance plus the breaking distance). With all locations having a 30 mph speed limit, the required stopping sight distance (assuming a 2.5 sec reaction time) is 47 m. Therefore, the benchmark points on the two Buffalo locations were set about 50 m before the crosswalk. For the Hamburg location, a benchmark of about half of this distance (about 22 m) was selected for two reasons. First many of the videos run from nearly 25-30 m before the crosswalk, and second, there were no easily identifiable landmarks in the 50 m range. The difference in benchmark positions was tediously addressed through the use of panel effects (random and fixed effects), and/or through the use of dummy indicator variables for the Hamburg location. Figure 3.8 illustrates the benchmark and crosswalk points in the HVC 6 (Oak/Eagle, Buffalo) location, presented as a snapshot of the forward facing video data.

Figure 3.8. HVC 6 Oak/Eagle: Forward facing video with benchmark and HVC points.

3.6. Measures of Effectiveness

Since there were no pedestrian-motor vehicle crashes in the Erie County SHRP2 NDS database and the data contained very few pedestrian-motor vehicle conflicts, the analysis of HVC effectiveness was based on surrogate measures. In road safety analysis, crash data is not always sufficient due to small sample sizes and lack of details about driver crash avoidance behavior. Past research (Tarko et al, 2009; Moreno & García, 2013) has shown that crash surrogates related to crashes can be used to capture the effect of a safety treatment without the occurrence of accidents. Crash and conflict related measures such as time-to-collision, crash potential index, and crash-to-surrogate ratio have been used as effective safety measures. Analysis of the speed profile was found to be another useful method to evaluate safety treatments in the absense of crash data. For this project, vehicle acceleration, speed, and gas pedal position during crosswalk traversals were used to capture changes in driving behavior before and after HVC installation. These driver behavior measures can impact pedestrian safety at crosswalks, and are therefore good surrogate safety measures.

Summary statistics of the surrogate safety measures at the benchmark point and crosswalk location are shown in Table 3.3, along with the change between the two points. The change between the two points of the three measures were used as dependent variables for statistical modeling. All surrogate measures shown in the table were used for the before-after analysis of the HVCs that were installed during the study.

Table 3.3. Summary statistics of surrogate measures.

Description	Mean	Standard Deviation	Minimum	Maximum
Speed at benchmark *(km/h)*	47.473	10.840	4.634	81.786
Speed at crosswalk *(km/h)*	48.720	11.443	5.647	81.927
Difference in speed from benchmark to crosswalk *(km/h)*	1.414	5.515	-30.906	24.691
Acceleration at benchmark *(g)*	0.0135	0.0599	-0.602	0.549
Acceleration at crosswalk *(g)*	0.0120	0.0449	-0.243	0.248
Difference in acceleration from benchmark to crosswalk *(g)*	-0.0015	0.0673	-0.504	0.769
Gas pedal position at benchmark	12.613	16.206	0	100
Gas pedal position at crosswalk	11.689	13.051	0	100
Difference in gas pedal position from benchmark to crosswalk	-0.923	16.453	-100	83.859

The values of speed (in km/h) ranged from very slow to much faster than the speed limit of 30 mph (48.3 km/h), however the average speed at both the benchmark and crosswalk was observed to be around the speed limit. The average change between the two points was an increase in speed, but both slowing and speeding up were observed.

Acceleration is measured in gravitational units (g), with a negative value signifying that the vehicle is decelerating. A negative change in acceleration between the benchmark and crosswalk points indicates that either the vehicle's acceleration decreased, or the deceleration increased. This is a change that would increase pedestrian safety. On the other hand, a positive change means that the vehicle's acceleration increased or the deceleration decreased. The gas pedal position is highly correlated with acceleration, as pressing on the gas pedal increases acceleration and releasing the gas pedal causes the vehicle to decelerate. It is measured from the vehicle network as the position of the pedal and normalized using the manufacturer specifications with a range of 0 to 100. Therefore, the potential range of difference from the benchmark to crosswalk is +/- 100.

3.7. Hypothesis Testing

The analysis included hypothesis tests for the locations (i.e., the two Buffalo crosswalks) with available data before and after the HVC installation. One-tail hypothesis *t*-tests were conducted to test whether there exist statistical differences in terms of reduction in speed, acceleration, and gas pedal position at the benchmark point, the HVC location, and between the two. The null hypothesis tested was that the average value of the measure was not different after the HVC installation from the value before. The test statistic for the hypothesis can be calculated as:

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{{s_1}^2}{n_1} + \frac{{s_2}^2}{n_2}}} \tag{1}$$

where: $\bar{X}$ is the mean value, s^2 is the variance, n is the number of observations, and $(\mu_1 - \mu_2)$ is the difference between the mean values in the null hypothesis, which is zero in this case. The subscripts of 1 and 2 represent the before and after conditions (Washington et al, 2011). Assuming a normal distribution, the significance of the *t*-value can be evaluated using the standard student's t distribution.

The above equation was used for each of the nine surrogate measures of effectiveness (i.e., speed, acceleration, and gas pedal position at the benchmark, crosswalk, and the change between the two points). The HVCs at the Buffalo locations were installed in two stages, first the crosswalk striping then the pedestrian crossing signs at the crosswalk and in advance. Therefore, three sets of hypothesis tests were conducted to analyze the impact of the striping, the pedestrian crossing signs, and the combination of the two. That is, the data was divided by the presence of the HVC striping, then by the presence of the pedestrian crossing signs, then finally by the presence of both versus the absence of both.

3.8. Statistical Modeling

Random parameter linear regression models were estimated, to determine the driver's reactions prior to crossing the crosswalk. The models used as dependent variables the change in speed, acceleration, and gas pedal position between the benchmark and crosswalk points. To that end, the standard linear regression model is given by:

$$Y_i = \beta_0 + \boldsymbol{\beta}_i \boldsymbol{X}_{in} + \varepsilon_{in} \tag{2}$$

where: Y_i is the dependent variable; β_0 is a constant term; $\boldsymbol{\beta}_i$ is the coefficient of explanatory variable $\mathbf{X}_{in}$ for observation *n*; and ε_{in} is the error term (Washington et al, 2011). Subscripts *i* and *n* represent the variable and observation, respectively.

The linear regression model in equation 1 assumes the effect of each explanatory variable in the **X** vector is the same for each trip observation. This means, for example, that the effect of snowy weather on a driver's speed is the same across all observations and individuals. In reality, there may be other unobserved factors that affect a driver's speed in snowy weather such as their confidence and familiarity with driving in poor conditions or the vehicle's tire traction. These influences of unobserved heterogeneity can be accounted for by incorporating random parameters which allow for the effect of each explanatory parameter to vary across observations. Random parameter linear regression models assume that the parameters vary according to a specified distribution. A normal distribution was found to provide the best fit for the models in this analysis.

Random parameters are introduced with $f(\boldsymbol{\beta}_i \mid \boldsymbol{\varphi})$, where $\boldsymbol{\varphi}$ is a vector of parameters of the density function (mean and variance). The resulting outcome probabilities are (see Anastasopoulos and Mannering, 2011):

$$P_n(i \mid \boldsymbol{\varphi}) = \int \frac{e^{\boldsymbol{\beta}_i \mathbf{X}_{in}}}{\sum_{\forall I} e^{\boldsymbol{\beta}_I \mathbf{X}_{In}}} f(\boldsymbol{\beta}_i \mid \boldsymbol{\varphi})\, d\boldsymbol{\beta}_i \tag{3}$$

where, $P_n(i \mid \boldsymbol{\varphi})$ is the outcome probability conditional on $f(\boldsymbol{\beta}_i \mid \boldsymbol{\varphi})$. For model estimation, $\boldsymbol{\beta}_i$ can account for variations of the effect of $\mathbf{X}$ on outcome probabilities, with the density function $f(\boldsymbol{\beta}_i \mid \boldsymbol{\varphi})$ used to determine $\boldsymbol{\beta}_i$. Mixed logit probabilities are then a weighted average for different values of $\boldsymbol{\beta}_i$ across drivers where some elements of the vector $\boldsymbol{\beta}_i$ may be fixed and some may be randomly distributed. Estimation of the random parameters multinomial logit model shown in Equation 2 is undertaken using simulated maximum likelihood approaches, and 200 Halton draws (Bhat, 2003). For the functional form of the parameter density functions, consideration was given to normal, lognormal, triangular, uniform and Weibull distributions, with the normal distribution consistently providing the best overall statistical fit. This feature of random parameters is important, as they can either capture the variable effect of a specific parameter on the dependent variable, or more importantly the effect of other unobserved factors.

The data used in this study is panel data, meaning that it is based on a cross-section of individuals observed over time. In addition to effects varying across observations, they may vary between the participants. There are unobserved human factors for each driver that causes their driving behavior to be different from others. These random effects are considered by incorporating the analysis of panel data. Since the number of trip observations was different for each driver, the data is an unbalanced panel of 62 individuals. The random parameter linear regression models that were estimated allow the effect of the parameters to vary by driver.

Other types of statistical models may also prove useful to evaluate the effectiveness of high-visibility crosswalks (HVCs), but random parameter linear regression was determined to be the best for this analysis. The continuous independent variables of change in the surrogate measures (i.e., speed, acceleration, and gas pedal position) can take negative values, therefore linear regression should be used. The analysis takes into account what factors affect the change in surrogate measures and can be used to show whether the HVC has a significant effect. Other model types, such as binary logit, would be helpful to evaluate whether the surrogate measures change over a specified threshold. However, this would require discretizing the data and valuable information would be lost. Whether there is a change or not is evaluated through hypothesis testing, but other factors are not taken into account as in statistical modeling.

3.9. Descriptive Statistics

The trip dataset used for analysis includes 1,078 trips made by a representative random sample of 62 participants that range in age from 16 to 84 years. The distribution of age by gender is depicted in Figure 3.9. Half of the participants were aged 34 or younger, however they only made approximately 36 percent of the trips across the HVC locations.

To capture the effect of familiarity with the location, the sample includes drivers who traversed the HVC locations many times as well as those who only made a few traversals. Figure 3.10 shows that 55 percent of trips were made by participants who traversed the same HVC location more than 50 times in the study. However, these trips were made by only 10 percent of the drivers, as shown in Figure 3.11. On the other hand, 70 percent of drivers traversed the same location less than 10 times and made up only 11 percent of the trips. This shows that the data selected has a distribution between drivers who were very familiar with the area and those that only traversed the HVC locations a few times.

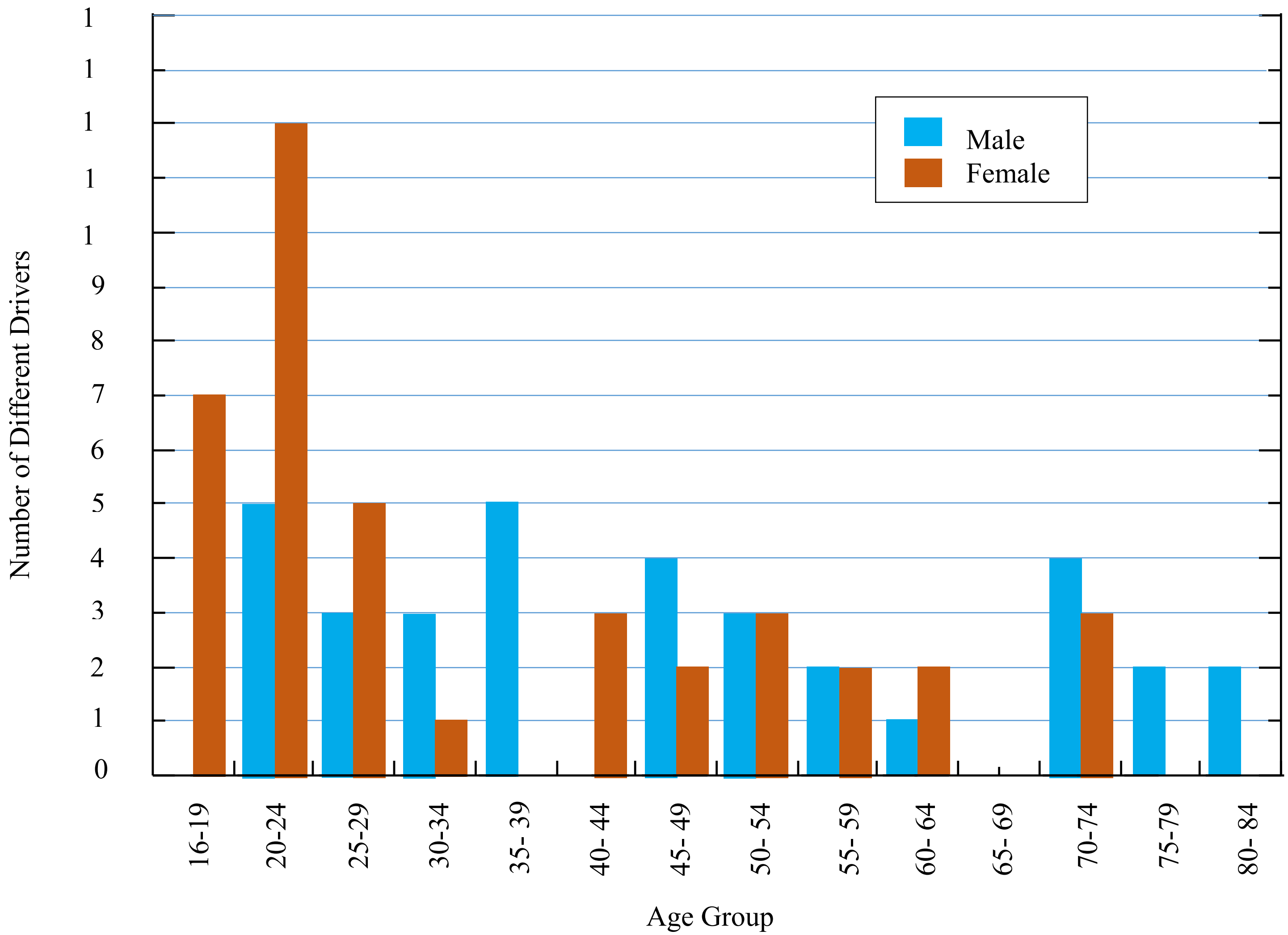

Figure 3.9. Gender and ages of drivers in the trip analysis dataset.

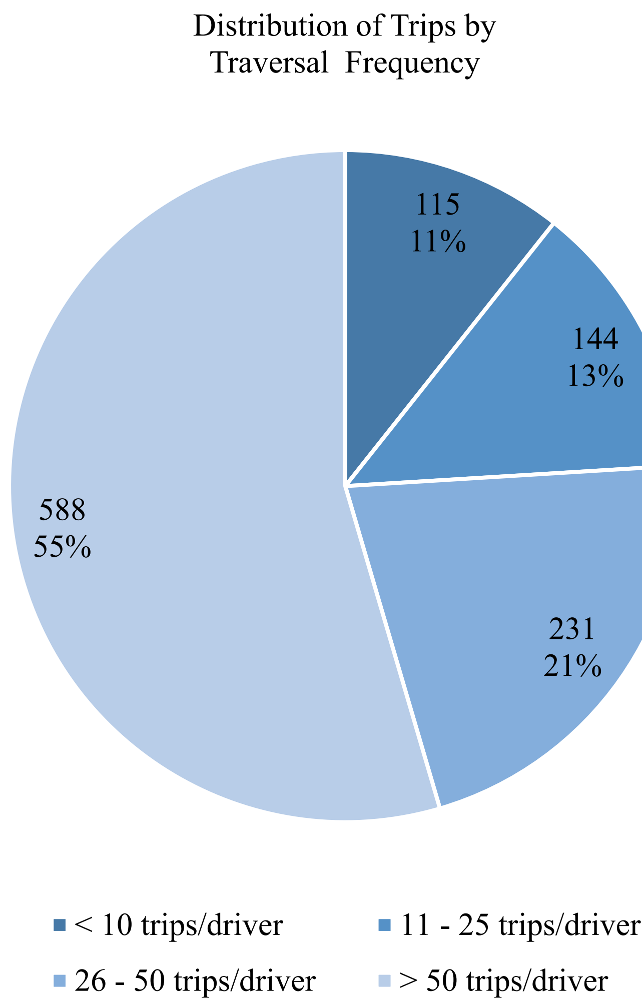

Figure 3.10. Distribution of trips by HVC traversal frequency of the driver during the study period.

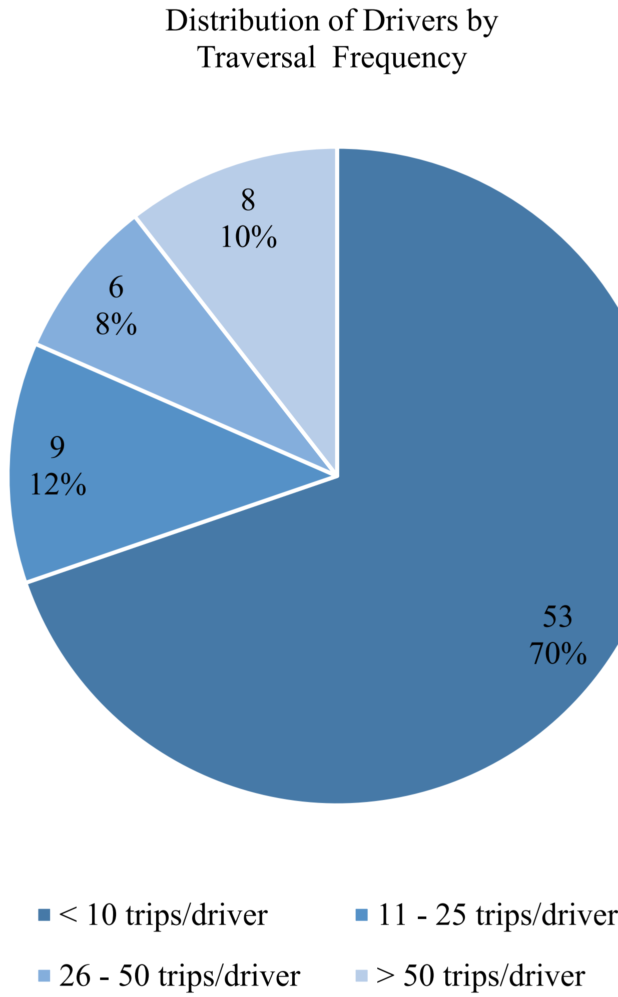

Figure 3.11. Distribution of drivers by HVC traversal frequency during the study period.

For use in statistical modeling, some binary indicator variables were defined from the available data. These include parameters such as the presence of pedestrians to indicate how drivers react when there are pedestrians in the vicinity of the crosswalk, and the presence of both the HVC striping and crossing sign. Variables indicating the driver's characteristics were also used, for example age under 30 years old (approximately 47 percent of trips) or frequent traveler over 50 trips. Environmental factors including the season, weather, and time of day were also considered. The distribution of trips by these factors are shown in Figure 3.12,Figure 3.13, and Figure 3.14. Two indicator variables that were used for these influences are winter trip and AM peak hour of 6-9am (22 and 25 percent of observations, respectively).

Another factor that was considered is how the driver was driving prior to the HVC location (i.e., at the benchmark location). Indicator variables were created based on the speed at the benchmark point to show if the speed was higher or lower than the speed limit of 30 mph by 5 mph (8.05 km/h) or more. The percentage of trips within each category (i.e., slow, fast, and near limit) are shown in Figure 3.15.

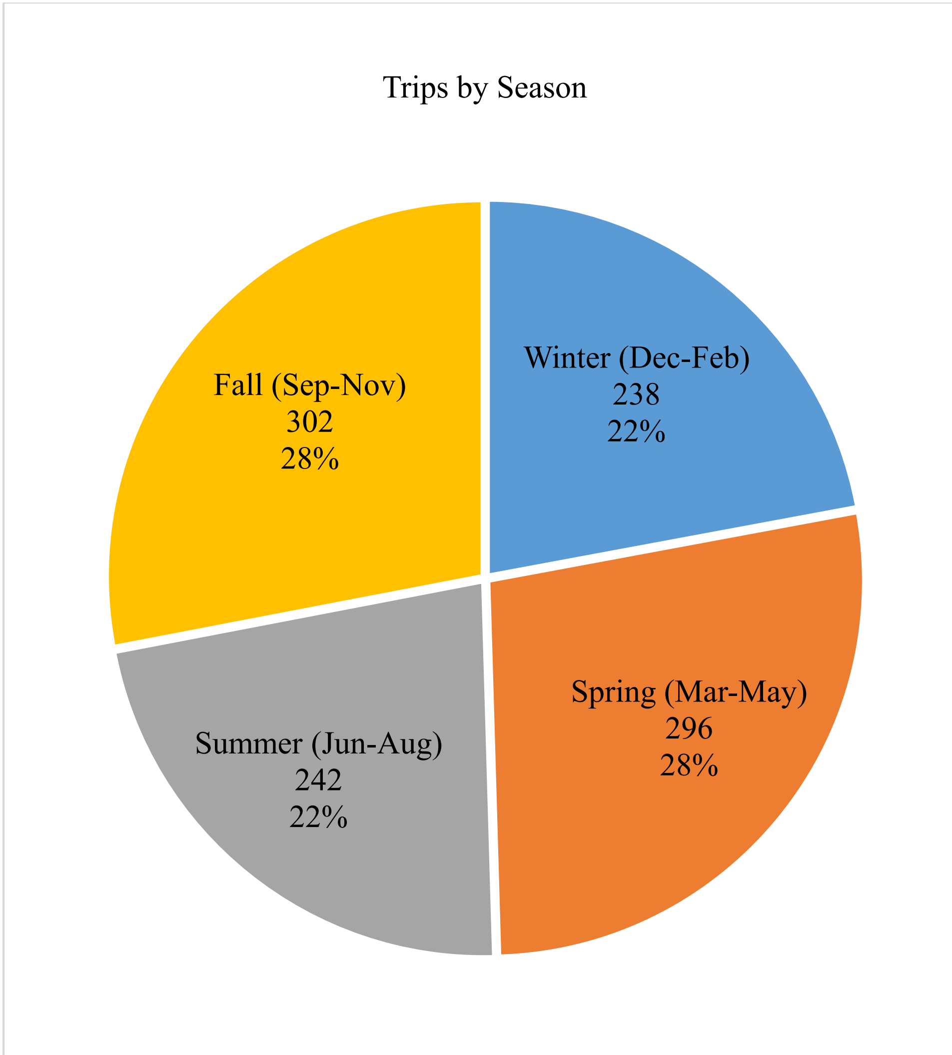

Figure 3.12. Percent of trips in each season of the year.

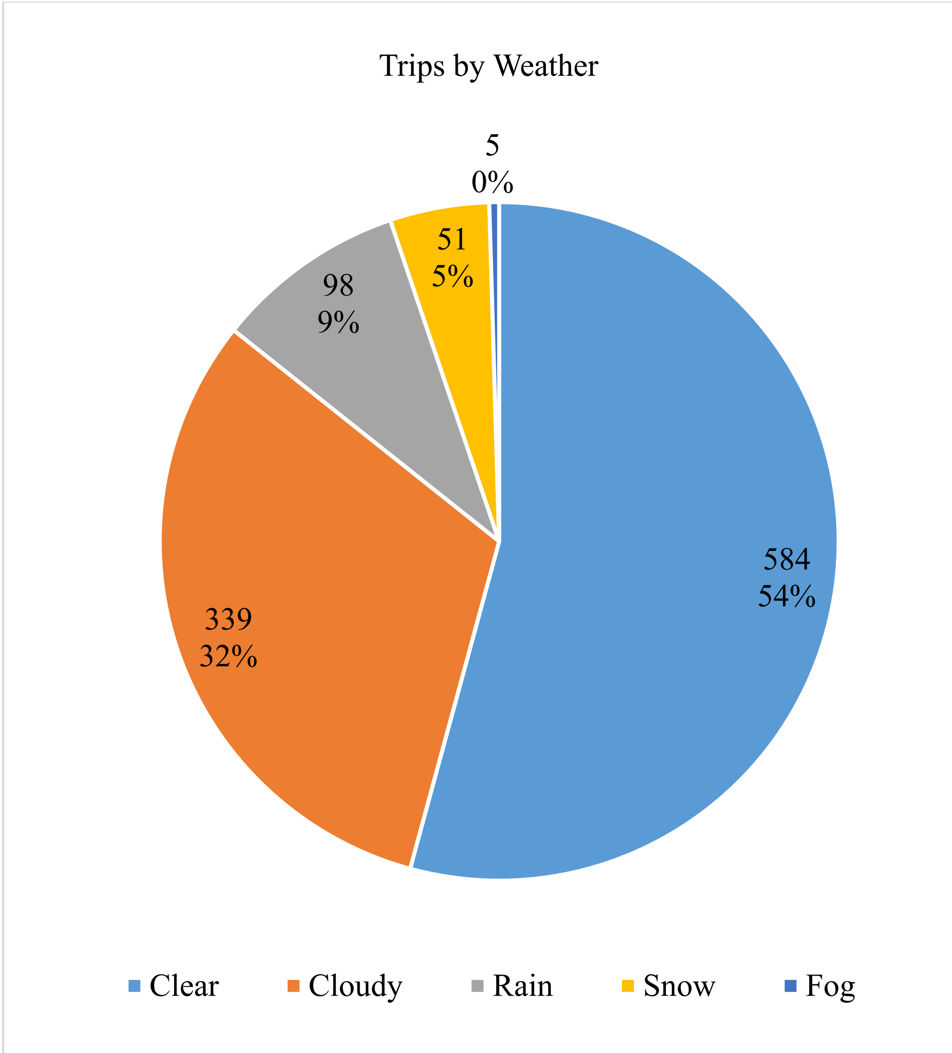

Figure 3.13. Trip percentage by weather at time of trip.

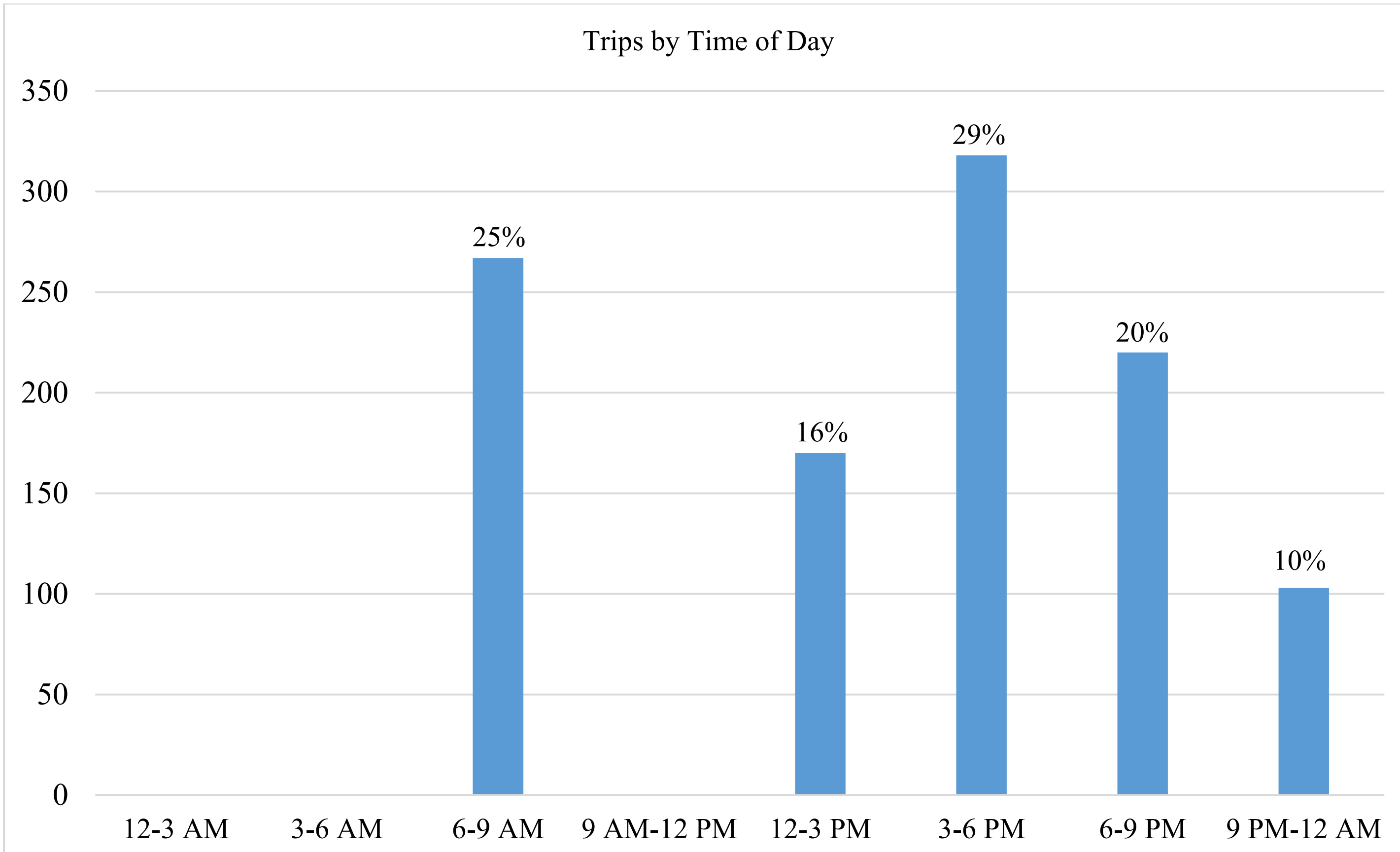

Figure 3.14. Distribution of trips throughout the day, divided into 3-hour time bins.

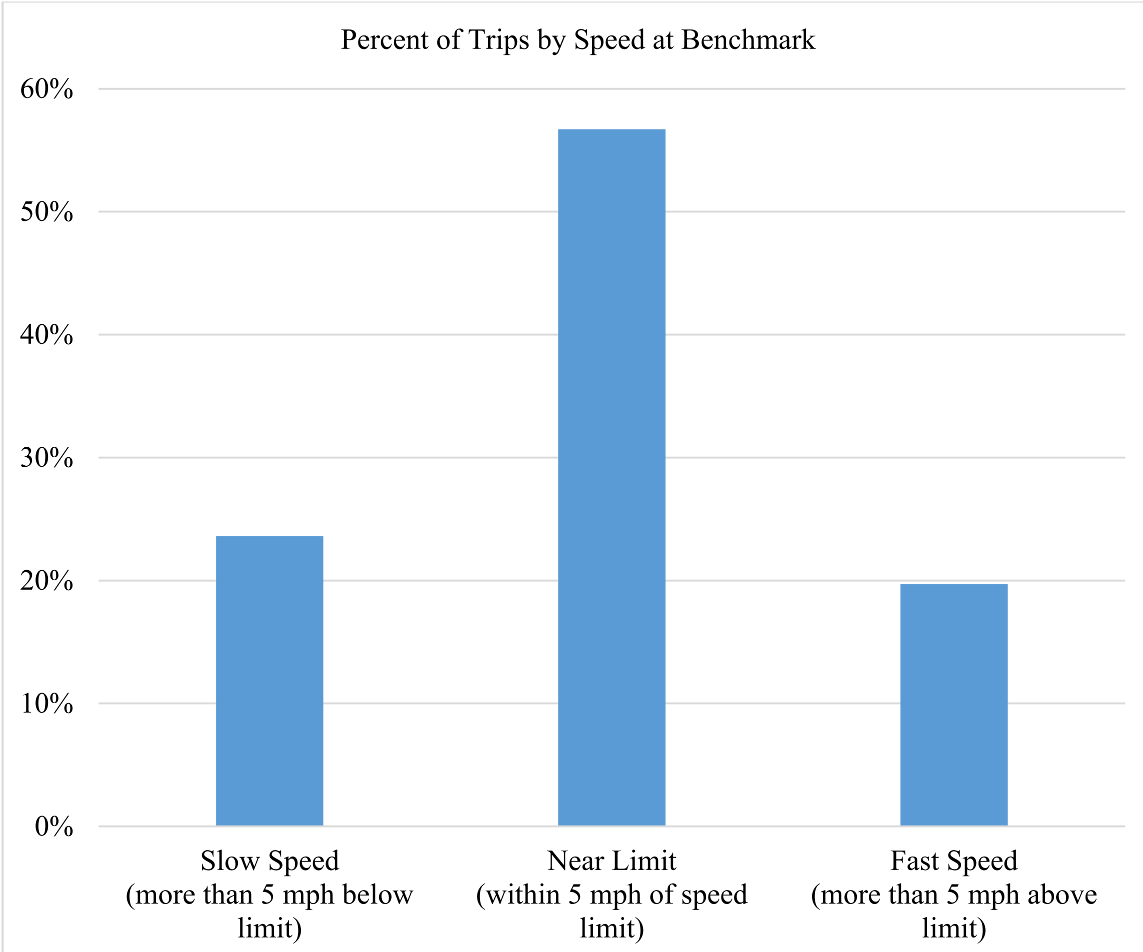

Figure 3.15. Trips distributed by speed range at benchmark location.

Other variables that were created from the data include the indication of a leading vehicle (43 percent of trips) and the presence of obstructing vehicles (25 percent with 3 or more, and 12 percent with 4 or more) that could be blocking the view of the HVC from the driver. Another visibility-related variable, the windshield condition, was taken into consideration. Approximately 13 percent of trips had a windshield condition that was classified as very poor through video observation. In addition, binary indicator variables for the HVC location were used to capture other unobserved effects related to the location such as number of lanes, lane width, and side clearance.

3.10. Summary

The data analysis began with selecting HVC locations in the Erie County SHRP2 study site, and a representative random sample of 1,078 trips through those locations. The various NDS data sources obtained for the trips (i.e., forward facing videos, time series data, and driver and vehicle information) were processed and combined into a manageable dataset for analysis. This required the selection of a benchmark point prior to the crosswalk location to represent where the drivers can see and react to the HVC.

Since there were no observed pedestrian – motor vehicle accidents, crash surrogate measures of speed, acceleration, and gas pedal position were used to analyze the HVC effectiveness. This included the values of the three measures at the benchmark and the crosswalk, as well as the change between the two locations. For the locations with data before and after HVC installation, hypothesis tests were performed to determine if the HVC striping and pedestrian crossing signs had a significant effect on the surrogate measures. Random parameters linear regression models were estimated for the difference in the surrogate measures between the two points. Unlike hypothesis testing, statistical modeling takes into account the effect of other factors. This exploratory analysis shows that NDS data is offers further insight into actual driving behavior compared to methods used in the past. Also, it will show whether HVCs impact driving behavior to improve pedestrian safety at unsignalized crosswalks.

CHAPTER 4. RESULTS AND DISCUSSION

4.1. Introduction

The effectiveness of high-visibility crosswalks (HVCs) in improving pedestrian safety was evaluated through hypothesis tests of before and after data, as well as random parameters linear regression modelling to account for other factors. The SHRP2 naturalistic driving study (NDS) data that was used provides detailed information of actual driving behavior at the three crosswalk locations in Buffalo and Hamburg, New York. The 62 participants' driving was recorded for 1- or 2-year time periods between 2011 and 2013, so many drivers traversed the same location multiple times. This allows the analysis to account for familiarity with the area before and after HVC installation. Three surrogate measures of effectiveness were analyzed: speed, acceleration, and gas pedal position. The results are presented and discussed in the subsequent sections.

4.2. Hypothesis Tests

For the locations with traversal data before and after the HVC installation, one-tail hypothesis *t*-tests were conducted to test whether there was a statistical change in speed, acceleration, and gas pedal position at the benchmark point, HVC location, and between the two. The HVCs at the Buffalo locations included crosswalk striping and pedestrian crossing signs at the crosswalk and in advance, with the striping installed prior to the two sets of signs. Before-after tests were performed for both measures to analyze the effectiveness of each separately. In addition, hypothesis tests were conducted to determine if the surrogate measures were changed from before both to after both to test the effect together. The average before and after values with corresponding *t*-tests are shown in Table 4.1.

Table 4.1. Average speed, acceleration, and gas pedal position, before and after HVC striping and pedestrian sign installation, and corresponding *t*-tests.

	HVC Striping Installation			Pedestrian Sign Installation			Both Striping and Ped. Sign Installation		
Variable	Before	After	*t*-score	Before	After	*t*-score	Before	After	*t*-score
Avg. Speed at Benchmark *(km/h)*	52.84	50.06	4.711*	52.06	50.03	3.112*	52.84	50.03	4.091*
Avg. Speed at HVC *(km/h)*	54.63	51.50	4.786*	53.70	51.58	2.908*	54.63	51.58	3.987*
Avg. Speed Difference Between Benchmark and HVC *(km/h)*	1.79	1.44	0.842	1.64	1.54	0.213	1.79	1.54	0.513
Avg. Acceleration at Benchmark *(g)*	0.018	0.015	0.635	0.016	0.017	-0.084	0.018	0.017	0.217
Avg. Acceleration at HVC *(g)*	0.017	0.002	4.575*	0.014	0.000	3.938*	0.017	0.000	4.050*
Avg. Acceleration Difference Between Benchmark and HVC *(g)*	-0.001	-0.013	2.883*	-0.002	-0.017	3.064*	-0.001	-0.017	3.234*
Avg. Gas Pedal Position at Benchmark	14.54	13.25	0.971	14.13	13.33	0.619	14.54	13.33	0.844
Avg. Gas Pedal Position at HVC	12.63	11.62	0.941	12.24	11.82	0.382	12.63	11.82	0.666
Avg. Gas Pedal Position Difference Between Benchmark and HVC	-1.91	-1.63	-0.210	-1.89	-1.50	-0.289	-1.91	-1.50	-0.276

Asterisks denote statistically different values at the 0.95 level of confidence (corresponding *t*-score is 1.645)

The results of hypothesis tests indicate that there was a statistically significant (at 0.95 level of confidence) reduction in speed at the benchmark and at the crosswalk location after the HVC striping, pedestrian signs, and after both were installed. There was also a statistically significant decrease in acceleration at the crosswalk after the striping, signs, and both. The difference in acceleration from the benchmark to the HVC was found to statistically significantly decrease after the installations, indicating that drivers were more likely to decelerate before the HVC location after the striping and pedestrian signs were installed. For most variables, the HVC striping had the most significant impact (highest *t*-score) compared to the pedestrian crossing sign and both measures together. However, this result alone does not indicate that crosswalk striping should be installed without some type of sign to warn drivers that there could be pedestrians trying to cross the roadway, as the pedestrian crossing signage was also significant.

4.3. Linear Regression

The first random parameters linear regression model uses as a dependent variable the change in speed from the benchmark point to the HVC location. A variety of factors was found to be significant, including the presence of pedestrians, the presence of the HVC and pedestrian sign, and obstructing vehicles. Table 4.2 and Table 4.3 show the descriptive statistics for all explanatory variables and the estimation results of the regression model, respectively. Since linear regression was used, the coefficients linearly correspond to the effect of the variable on the speed change.

The results show that the combination of the HVC and pedestrian sign is related to a decrease in speed of 1.07 km/h from the benchmark to the crosswalk. Also, the presence of a pedestrian near the crosswalk caused a -0.90 km/h change in speed. If there was a leading vehicle or at least 4 vehicles obstructing the view of the crosswalk, the speed decreased. The leading vehicle indicator is a random parameter, therefore based on the normal distribution of the coefficient the effect is negative about 90 percent of the time. This shows that drivers were more likely to slow down before the crosswalk if there were other vehicles blocking the view of potential pedestrians. It could also be capturing the effect of surrounding vehicles slowing in traffic. The speed change also decreased if the vehicle exceeded the speed limit at the benchmark point by at least 5 mph (i.e., 8.05 km/h), the speed was likely to decrease between the benchmark and the HVC (0.72 probability of negative effect). This indicates that the drivers tended to slow down before the crosswalk if they had been speeding previously.

Table 4.2. Descriptive statistics of explanatory variables for speed change model.

Variable Description	Mean	Std. Dev.	Min	Max
Dependent Variable: Difference in speed (from GPS) from benchmark to HVC *(km/h)*	1.414	5.515	-30.906	24.691
Pedestrian indicator (1 if pedestrian is present near the HVC, 0 otherwise)	0.146	0.353	0	1
HVC and pedestrian sign indicator (1 if both are present, 0 otherwise)	0.494	0.500	0	1
Leading vehicle indicator (1 if leading vehicle is present, 0 otherwise)	0.429	0.495	0	1
Obstructing vehicle indicator (1 if there are 4 or more vehicles obstructing the view to the crosswalk, 0 otherwise)	0.116	0.320	0	1
Time indicator (1 if traversal occurred between 6 am and 9 am, 0 otherwise)	0.248	0.432	0	1
Driver's age indicator (1 if less than 30 years old, 0 otherwise)	0.429	0.495	0	1
Frequent traveler indicator (1 if driver traversed more than 50 times the same location, 0 otherwise)	0.545	0.498	0	1
Speeding indicator (1 if the vehicle exceeds the speed limit at the benchmark point by 5 mph - 8.05 km/h - or more, 0 otherwise)	0.197	0.398	0	1
Location indicator (1 if Hamburg location, 0 otherwise)	0.256	0.437	0	1
Speed at benchmark *(km/h)*	47.473	10.840	4.634	81.786

Table 4.3. Estimation results of regression model for speed change.

Variable Description	Coefficient	*t*-statistic	*p*-value
Constant	9.454	3.55***	0.0004
Standard deviation of parameter density function	*1.798*	10.62***	*0.0000*
Pedestrian indicator (1 if pedestrian is present near the HVC, 0 otherwise)	*-0.896*	-1.73*	*0.0843*
HVC and pedestrian sign indicator (1 if both are present, 0 otherwise)	-1.071	-2.11**	0.0345
Leading vehicle indicator (1 if leading vehicle is present, 0 otherwise)	-1.555	-3.02***	0.0025
Standard deviation of parameter density function	*1.211*	5.14***	*0.0000*
Obstructing vehicle indicator (1 if there are 4 or more vehicles obstructing the view to the crosswalk, 0 otherwise)	-1.258	-2.04**	0.0416
Time indicator (1 if traversal occurred between 6 am and 9 am, 0 otherwise)	1.187	2.75***	0.0060
Driver's age indicator (1 if less than 30 years old, 0 otherwise)	1.348	3.24***	0.0012
Standard deviation of parameter density function	*1.591*	5.23***	*0.0000*
Frequent traveler indicator (1 if driver traversed more than 50 times the same location, 0 otherwise)	1.860	6.9***	0.0000
Speeding indicator (1 if the vehicle exceeds the speed limit at the benchmark point by 5 mph - 8.05 km/h - or more, 0 otherwise)	-2.003	-2.25**	0.0245
Standard deviation of parameter density function	*3.435*	6.78***	*0.0000*
Location indicator (1 if Hamburg location, 0 otherwise)	-2.938	-3.91***	0.0001
Speed at benchmark predictor *(km/h)*	-0.152	-2.88***	0.0040
Variance parameter, sigma	4.198	85.36***	0
Number of individuals / Number of observations	60 / 982		
Log likelihood function	-2811.535		
Restricted log likelihood	-3069.256		

Note: ***, **, * indicate significance at 1%, 5%, 10% level

Speed increased between the benchmark and crosswalk if the trip was during the morning hours of 6 to 9 am, as well as if the driver was a frequent traveler through the intersection and was familiar with the area. If the driver is under 30 years old, there was a 0.2 probability that the speed change between the benchmark and crosswalk would increase. An indicator variable for the Hamburg location was found to be statistically significant, and captures the effects of roadway characteristics and other factors.

A linear regression model for difference in acceleration from the benchmark to the crosswalk was also estimated. Unlike the speed difference model, no variables were found to be significant as random parameters. As in the speed change model, the presence of the HVC and pedestrian sign was found to be an explanatory variable. The model accounted for other factors including visibility, speed, and location differences. The descriptive statistics for all parameters are shown in Table 4.4, and the model results are in Table 4.5.

Table 4.4. Descriptive statistics of explanatory variables for acceleration change model.

Variable Description	Mean	Std. Dev.	Min	Max
Dependent Variable: Difference in acceleration from benchmark to HVC *(g)*	-0.00148	0.0673	-0.5038	0.7688
HVC and pedestrian sign indicator (1 if both are present, 0 otherwise)	0.4935	0.5002	0	1
Obstructing vehicle indicator (1 if there are 3 or more vehicles obstructing the view to the crosswalk, 0 otherwise)	0.2449	0.4302	0	1
Windshield condition indicator (1 if visibility through windshield is very poor, 0 otherwise)	0.1299	0.3363	0	1
Slow speed indicator (1 if the vehicle is below the speed limit at the benchmark point by 5 mph - 8.05 km/h - or more, 0 otherwise)	0.2356	0.4246	0	1
Location indicator (1 if Buffalo-Elm/Eagle location, 0 otherwise)	0.4397	0.4966	0	1
Acceleration at benchmark *(g)*	0.0135	0.0600	-0.6016	0.5491

Table 4.5. Estimation results of regression model for acceleration change.

Variable Description	Coefficient	*t*-statistic	*p*-value
Constant	0.0139	3.41***	0.0006
HVC and pedestrian sign indicator (1 if both are present, 0 otherwise)	-0.0079	-1.79*	0.0736
Obstructing vehicle indicator (1 if there are 3 or more vehicles obstructing the view to the crosswalk, 0 otherwise)	-0.0090	-1.83*	0.0676
Windshield condition indicator (1 if visibility through windshield is very poor, 0 otherwise)	-0.0147	-2.40**	0.0162
Slow speed indicator (1 if the vehicle is below the speed limit at the benchmark point by 5 mph - 8.05 km/h - or more, 0 otherwise)	0.0182	3.44***	0.0006
Location indicator (1 if Buffalo-Elm/Eagle location, 0 otherwise)	-0.0101	-2.28**	0.0223
Acceleration at benchmark predictor *(g)*	-0.6602	-10.64***	0.0000
Number of individuals / Number of observations	62 / 987		
Log likelihood function	1335.660		
Restricted log likelihood	1256.710		

Note: ***, **, * indicate significance at 1%, 5%, 10% level

Table 4.5 shows that the presence of both the HVC and pedestrian sign decreased the change in acceleration (by -0.008 g), or increased the amount of deceleration between the benchmark and crosswalk. Deceleration also increased if visibility was limited due to either obstructing vehicles or very poor windshield condition. On the other hand, acceleration change increased if the vehicle was traveling slower than the speed limit at the benchmark point by at least 5 mph (i.e., 8.05 km/h) indicating that drivers tend to accelerate leading up to the intersection if they had been driving slowly. A location indicator was also included to capture the effects related to the roadway characteristics and other factors specific to the Elm/Eagle location.

The final linear regression model used change in gas pedal position between the benchmark point and HVC location as the dependent variable. No random parameters were found to be significant in the model. Explanatory variables that were found to be significant are shown in Table 4.6, and Table 4.7 shows the results of model estimation.

Table 4.6. Descriptive statistics of explanatory variables for gas pedal position change model.

Variable Description	Mean	Std. Dev.	Min	Max
Dependent Variable: Difference in gas pedal position from benchmark to HVC	-0.923	16.453	-100	83.859
Season indicator (1 if traversal occurred between Dec. and Feb., 0 otherwise)	0.221	0.415	0	1
HVC and slow speed indicator (1 if the vehicle is below the speed limit at the benchmark point by 5 mph - 8.05 km/h - or more and HVC is not installed, 0 otherwise)	0.019*	0.135	0	1
Gas pedal position at benchmark	12.613	16.206	0	100

Note: This variable has the problem of low variability and warrants further investigation, but serves as an indication for the effect of the HVC on gas pedal position change.

Table 4.7. Estimation results of regression model for gas pedal position change.

Variable Description	Coefficient	*t*-statistic	*p*-value
Constant	3.878	4.00***	0.0001
Season indicator (1 if traversal occurred between Dec. and Feb., 0 otherwise)	2.114	1.69*	0.0913
HVC and slow speed indicator (1 if the vehicle is below the speed limit at the benchmark point by 5 mph - 8.05 km/h - or more and HVC is not installed, 0 otherwise)	7.288	1.72*	0.0856
Gas pedal position at benchmark predictor	-0.420	-7.52***	0.0000
Number of individuals / Number of observations	62 / 910		
Log likelihood function	-3824.861		
Restricted log likelihood	-3855.322		

Note: ***, **, * indicate significance at 1%, 5%, 10% level

As shown in the results, a small number of parameters were found to be significant and provided the best statistical fit of the model. If the vehicle was traveling below the speed limit at the benchmark by 5 mph (i.e., 8.05 km/h) or more, and the HVC was not installed, the gas pedal position increased. This could be analyzed by considering the opposite case, when the HVC is installed, and there would be a negative change in gas pedal position. This variable warrants further inspection because it has low variability and is potentially accounting for other effects, but it serves as an indication for the effect of the HVC on gas pedal position change. The season was also found to be an explanatory parameter, and gas pedal position change increased during the winter months.

4.4. Discussion

For the locations with available data before and after installation, the HVC and pedestrian crossing signs were found to decrease the average speed at the benchmark location, as well as the average speed and acceleration at the crosswalk location. After HVC installation, the acceleration change between the benchmark and crosswalk decreased. With decreased speed and acceleration, pedestrian safety at the crosswalk was increased as a result of the HVC installation.

Statistical modeling takes into effect other factors that hypothesis tests do not, and also showed that the presence of the HVC and pedestrian crossing signs decrease speed and acceleration change from the benchmark to the crosswalk. Other factors such as obstructions, visibility, time of day, season, and driver's traversal frequency were found to affect the surrogate measures of HVC effectiveness. To evaluate the statistical fit of the models, goodness-of-fit measures were calculated, shown in Table 4.8.

To avoid calculation issues due to small denominators, near-zero observed values (i.e., between +/- 1 km/h for speed change, +/- 0.02 g for acceleration change, and +/- 1 for gas pedal position change) were removed before calculating the accuracy measures. Based on the MAPE values, which account for the different sample sizes of the models, the speed change model has the best fit with a MAPE value of 0.864. Other measures cannot be accurately compared across the models due to the different scales in the actual values. Because of missing data, especially for gas pedal position change, the models do not predict the observed values as well as they could with a larger dataset. The 1,078 trip dataset used for this project is relatively small compared to all HVC traversals in the SHRP2 NDS data, but it is large enough to demonstrate the feasibility of using NDS data to evaluate the effectiveness of HVCs in improving pedestrian safety.

Table 4.8. Goodness-of-Fit Measures

Accuracy Measure	Linear Regression Model		
	Speed Change	Acceleration Change	Gas Pedal Change
Mean Absolute Percentage Error (MAPE)	0.864	0.889	1.125
Mean Error (ME)	0.023	0.001	-0.009
Mean Absolute Deviation (MAD)	3.232	0.056	12.019
Mean Squared Error (MSE)	21.813	0.006	330.893
Root Mean Squared Error (RMSE)	4.670	0.081	18.190
Standard Deviation of Errors (SDE)	4.674	0.081	18.203
Mean Percentage Error (MPE)	0.714	0.850	0.990

For the change in speed, acceleration, and gas pedal position, the respective variables at the benchmark point were used as explanatory parameters, which inevitably introduced endogeneity. This misspecification issue was treated by regressing the endogenous variables against all exogenous variables, and using their predictors for model estimation. Use of the values at the benchmark locations were found to improve the statistical fit of all three models because all other significant parameters were binary indicator variables.

4.5. Summary

The analysis shows that HVCs modify driving behavior and have the potential to improve pedestrian safety. Specifically, the HVC in combination with pedestrian crossing signs were related to lower speeds at the benchmark and crosswalk and lower acceleration at the crosswalk. Also, there was a greater decrease in speed and acceleration from the benchmark to the crosswalk position with the presence of the HVC. There is also a potential decrease in gas pedal position, but that relationship requires further analysis with a larger dataset. Potential for future work to further analyze HVC effectiveness using NDS data is discussed in the next chapter.

CHAPTER 5. SUMMARY AND CONCLUSION

5.1. Summary

Making roadways safer for pedestrians is an important goal in the United States and New York State. One widely employed strategy to accomplish this goal is the use of high-visibility crosswalk (HVC) markings. The effectiveness of safety countermeasures has been analyzed using various data types and methodologies. This project evaluated the effectiveness of HVCs to improve pedestrian safety at uncontrolled locations using the SHRP2 naturalistic driving study (NDS) data. NDS data offers a unique opportunity to analyze actual driving behavior over a period of time along with many other parameters. Three uncontrolled crosswalk locations in the Erie County, New York test site were selected for analysis. Two of the locations had HVCs installed during the study period, allowing for a before/after analysis. At the third location, only post HVC installation data was available. A representative random sample of 1,078 trips by 62 participants was selected for the study. For each trip, forward-facing video, time series data, and basic driver and vehicle information was processed and compiled into the dataset for analysis.

No pedestrian – motor vehicle crashes were observed, so crash surrogates (i.e., speed, acceleration, and gas pedal position) were used to evaluate driving behavior.

Random parameters linear regression models were estimated for the change in the surrogate measures between a benchmark point and the crosswalk while controlling for a variety of other factors. Hypothesis tests were conducted using data from the locations with observations before and after the HVC was installed. The key findings and potential for future work are discussed in the following sections. Overall, this work shows that HVCs have the potential to improve pedestrian safety and modify driving behavior, and that NDS data is useful for analyzing their effectiveness.

5.2. Key Findings

The effectiveness of high-visibility crosswalks (HVC) to improve pedestrian safety at uncontrolled locations was evaluated by analyzing the driving behavior of SHRP2 participants at three locations in the Erie County, New York test site. In the absence of pedestrian-motor vehicle crashes, speed, acceleration, and gas pedal position were used as surrogate safety measures. Table 5.1 shows the desirable effects of the surrogate measures and indicates which ones were proven through the analysis.

Table 5.1. Desirable effects of HVCs on surrogate measures with results.

Parameter	Desirable Effect for Pedestrian Safety	Proven
Speed at Benchmark *(km/h)*	Slower speed	✓
Speed at HVC *(km/h)*	Slower speed	✓
Speed Difference Between Benchmark and HVC *(km/h)*	Decrease (more slowing between benchmark and HVC)	✓
Acceleration at Benchmark *(g)*	Lower (less acceleration or more deceleration)	
Acceleration at HVC *(g)*	Lower (less acceleration or more deceleration)	✓
Acceleration Difference Between Benchmark and HVC *(g)*	Decrease (less acceleration or more deceleration between benchmark and HVC)	✓
Gas Pedal Position at Benchmark	Lower (less pressure on gas pedal)	
Gas Pedal Position at HVC	Lower (less pressure on gas pedal)	
Gas Pedal Position Difference Between Benchmark and HVC	Decrease (let up on gas pedal between benchmark and HVC)	✓*

Note: This result warrants further investigation, but indicates the potential effect of the HVC on gas pedal position change.

From the estimated statistical models, it was found that the HVC combined with the pedestrian crossing sign decreases the change in speed and acceleration between the benchmark and crosswalk points. Potentially, the HVC could also decrease gas pedal position, and this could be further analyzed with a larger dataset. In general, HVCs can modify driving behavior, which in turn increases pedestrian safety at uncontrolled locations. As an exploratory analysis this work also shows that naturalistic driving study data are useful for analyzing HVC effectiveness.

5.3. Directions for Future Research

This study was limited by the amount of data that were used. It takes many hours to process the videos and the time series data for all trips. An automated method to process videos would accelerate the process and expand the amount of data that could be analyzed. With more trip observations, statistical models could provide a better statistical fit to evaluate the effectiveness of HVCs. Specifically for surrogate measures and other parameters that have missing data in some trips due to the equipment, more available trip data would prove to be useful.

In addition to purely more data, many other locations should be analyzed both within the Erie County SHRP2 study site and in the other five sites to capture regional effects. Drivers in different areas have different driving behaviors and react to pedestrians differently. Also, policy makers in the other sites may enforce crosswalk laws differently and implement HVCs of various types and in different roadway locations. This analysis focused on one type of HVC installation (i.e., high-visibility striping and pedestrian crossing signage), but there are many other types of HVCs in the SHRP2 data. It would be beneficial to evaluate and compare the effectiveness of various HVC designs and signage in modifying driver behavior.

There is some data that is available in the SHRP2 NDS database that were not used for this exploratory analysis. Each participant in the study had medical and mental examinations, and answered survey questions about topics such as their personal driving behavior and their understanding of driving rules. Use of this data in statistical modeling would account for more specific driver characteristics. More specific roadway and pavement conditions could also be accounted for through the use of the Roadway Information Database (RID) that was compiled for each SHRP2 study site. This includes characteristics such as lane width, side clearance, and pavement condition which may also affect driving behavior.

With more trip data from other study sites and more HVC locations, there would be more observations of pedestrians crossing the roadway. The three sites that were used in this analysis do not have high pedestrian volumes, especially at night. Future research should further evaluate the effect that pedestrians have on driving behavior at uncontrolled HVCs. Separate models could be estimated for observations with pedestrians crossing and no pedestrians, to determine if the effectiveness of HVCs is different whether there are pedestrians present or not. In addition, separate models could be estimated for before and after HVC installation to compare the effects of other factors with and without the HVC presence.

The empirical analysis of HVC effectiveness performed in this study shows that HVCs can modify driving behavior in terms of speed, acceleration, and gas pedal position change. It is also shown that NDS data is useful evaluate driving behavior at uncontrolled HVC locations. The SHRP2 NDS data offers a unique opportunity to capture actual driver behavior. With the rich source of data, future research could expand on the analysis and further the understanding of the effectiveness of various types of HVCs in different locations.

REFERENCES

Anastasopoulos, I., Anastasopoulos, P.Ch., Agalianos, A., Sakellariadis, L., 2015. Simple method for real–time seismic damage assessment of bridges. *Soil Dynamics and Earthquake Engineering*, 78, 201-212.

Anastasopoulos, P.Ch., Islam, M., Volovski, M., Powell, J., Labi, S., 2011a. Comparative evaluation of public-private partnerships in roadway preservation. *Transportation Research Record*, 2235, 9-19.

Anastasopoulos, P.Ch., Haddock, J., Karlaftis, M., Mannering, F., 2012a. Analysis of urban travel times: Hazard-based approach to random parameters. *Transportation Research Record*, 2302, 121-129.

Anastasopoulos, P.Ch., Haddock, J., Peeta, S., 2014. Cost overrun in public-private partnerships: Toward sustainable highway maintenance and rehabilitation. *ASCE Journal of Construction Engineering and Management*, 140(6), 04014018.

Anastasopoulos, P.Ch., Karlaftis, M., Haddock, J. Mannering, F., 2012b. Household automobile and motorcycle ownership analyzed with random parameters bivariate ordered probit model. *Transportation Research Record*, 2279, 12-20.

Anastasopoulos, P.Ch., Labi, S., Bhargava, A., Mannering, F. 2012c. Empirical assessment of the likelihood and duration of highway project time delays. *ASCE Journal of Construction Engineering and Management,* 18(3), 390-398.

Anastasopoulos, P.Ch., Labi, S., Karlaftis, M., Mannering, F., 2011b. Exploratory State-level empirical assessment of pavement performance. *ASCE Journal of Infrastructure Systems*, 17(4), 200-215.

Anastasopoulos, P.Ch., Labi S., McCullouch, B., 2009. Analyzing duration and prolongation of performance-based contracts using hazard-based duration and zero-inflated random parameters Poisson models. *Transportation Research Record*, 2136, 11-19.

Anastasopoulos, P.Ch., Labi S., McCullouch, B., Karlaftis, M., Moavenzadeh, F., 2010a. Influence of highway project characteristics on contract type selection: Empirical assessment. *ASCE Journal of Infrastructure Systems*, 16(4), 323-333.

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

Anastasopoulos, P.Ch., Mannering, F.L, 2011. An empirical assessment of fixed and random parameter logit models using crash- and non-crash-specific injury data. *Accident Analysis and Prevention, 43*(3), 1140-1147.

Anastasopoulos, P.Ch., Mannering, F.L., 2014. Analysis of pavement overlay and replacement performance using random-parameters hazard-based duration models. *ASCE Journal of Infrastructure Systems*, 10.1061/(ASCE)IS.1943-555X.0000208 , 04014024

Anastasopoulos, P.Ch., Mannering, F.L., Haddock, J., 2012d. A random parameters seemingly unrelated equations approach to the post-rehabilitation performance of pavements. *ASCE Journal of Infrastructure Systems*, 18(3), 176-182.

Anastasopoulos, P.Ch., Mannering, F.L., Shankar, V. N., & Haddock, J. E, 2012e. A study of factors affecting highway accident rates using the random-parameters tobit model. *Accident Analysis and Prevention, 45*, 628-633.

Anastasopoulos, P.Ch., McCullouch, B., Gkritza, K., Mannering, F., Sinha, K.C., 2010b. Cost Savings Analysis of Performance-Based Contracts for Highway Maintenance Operations. *ASCE Journal of Infrastructure Systems*, 16(4), 251-263.

Anastasopoulos, P.Ch., Shankar, V., Haddock, J., Mannering, F., 2012f. A multivariate tobit analysis of highway accident-injury-severity rates. *Accident Analysis and Prevention*, 45, 110-119.

Anastasopoulos, P.Ch., Tarko, A., Mannering, F., 2008. Tobit analysis of vehicle accident rates on interstate highways. *Accident Analysis and Prevention*, 40(2), 768-775.

Anwaar, A., Anastasopoulos, P.Ch., Ong, G. Labi, S., Islam, M., 2012. Factors affecting highway safety, health care services, and motorization - An exploratory empirical analysis using aggregate data. *Journal of Transportation Safety and Security*, 4(2), 94-115.

Aziz, H. A., Ukkusuri, S. V., & Hasan, S, 2013. Exploring the determinants of pedestrian-vehicle crash severity in New York City. *Accident Analysis and Prevention, 50*, 1298-1309.

Bella, F., & Silvestri, M., 2015. Effects of safety measures on driver's speed behavior at pedestrian crossings. *Accident Analysis and Prevention, 83*, 111-124.

Bhat, C. R., 2003. Simulation estimation of mixed discrete choice models using randomized and scrambled Halton sequences. *Transportation Research Part B: Methodological, 37*(9), 837-855.

Campbell, K., 2012. The SHRP 2 Naturalistic Driving Study: Addressing Driver Performance and Behavior in Traffic Safety. *TR News September-October 2012: Blueprints to Improve Highway Safety*, 30-36.

FHWA, 2015. SHRP2 Industry Briefing [PowerPoint]. Obtained from http://www.fhwa.dot.gov/goshrp2/Resources

Fitzpatrick, K., Turner, S. M., Brewer, M., Carlson, P. J., Ullman, B., Trout, N. D., . . . Lord, D, 2006. *Improving Pedestrian Safety at Unsignalized Crossings.* Transportation Research Board: TCRP Report 112, NCHRP Report 562.

Gómez, R. A., Samuel, S., Romoser, M. R., Gerardino, L. R., Knodler, M., Collura, J., & Fisher, D. L., 2013. A driving simulator evaluation of road markings and symbolic signs on vehicle-pedestrian conflicts. *Road safety on four continents: 16th international conference.* Beijing, China.

Haleem, K., Alluri, P., & Gan, A., 2015. Analyzing pedestrian crash injury severity at signalized and non-signalized locations. *Accident Analysis & Prevention, 81*, 14-23.

Hedlund, J., 2015. SHRP 2 naturalistic driving study: The first data anslysis studies. *Journal of Safety Research, 54*, 3-4.

Herms, B. F., 1972. Pedestrian Crosswalk Study: Accidents in Painted and Unpainted Crosswalks. *Highway Research Record 406, Highway Research Board*, 1-14.

Jones, T. L., & Tomcheck, P., 2000. Pedestrian Accidents in Marked and Unmarked Crosswalks: A Quantitative Study. *ITE Journal, 70*(9), 42-46.

Knoblauch, R. L., & Raymond, P. D., 2000. *The Effect of Crosswalk Markins on Vehicle Speeds in Maryland, Virginia, and Arizona.* Report No. FHWA-RD-00-101, Federal Highway Administration, Washington, D.C.

Knoblauch, R. L., Nitzburg, M., & Seifert, R. F., 2001. *Pedestrian Crosswalk Case Studies: Sacramento, California; Richmond, Virginia; Buffalo, New York; Stillwater, Minnesota.* Report No. FHA-RD-00-103, Federal Highway Administration, Washington, D.C.

McGrane, A., & Mitman, M., 2013. *An Overview and Recommendations of High-Visibility Crosswalk Marking Styles.* University of North Carolina, Highway Safety Research Center, Pedestrian and Bicycle Information, Chapel Hill, NC.

Moreno, A. T., & García, A., 2013. Use of speed profile as surrogate measure: Effect of traffic calming devices on crosstown road safety performance. *Accident Analysis and Prevention, 61*, 23-32.

New York State, 2014. *New York State Highway Safety Annual Report FFY 2014.* Retrieved from http://www.nhtsa.gov/links/StateDocs/pages/AnnualReports.htm

NHTSA, 2013. *Traffic Safety Facts: Pedestrians.* Retrieved from http://www-nrd.nhtsa.dot.gov/Pubs/812124.pdf

Nitzburg, M., & Knoblauch, R. L., 2001. *An Evaluation of High-Visibility Crosswalk Treatments-Clearwater, Florida.* Report No. FHWA-RD-00-105, Federal Highway Administration, Washington, D.C.

Olszewski, P., Szagala, P., Wolański, M., & Zielińska, A., 2015. Pedestrian fatality risk in accidents at unsignalized zebra crosswalks in Poland. *Accident Analysis and Prevention, 84*, 83-91.

Russo, B.J., Savolainen, P.T., Schneider IV, W.H., Anastasopoulos, P.Ch., 2014. Comparison of factors affecting injury severity in angle collisions by fault status using a random parameters bivariate ordered probit model. *Analytic Methods in Accident Research, 2*, 21-29.

Samuel, S., Romoser, M. R., Gerardino, L. R., Hamid, M., Gómez, R. A., Knodler Jr, M. A., . . . Fisher, D., 2013. Effects of Advance Yield Markings and Symbolic Signs on Vehicle-Pedestrian Conflicts: Field Evaluation. *Transportation Research Record: Journal of the Transportation Research Board*, 139-146.

SHRP2. System and Installation. http://www.shrp2nds.us/DAS-installation.html

Tarko, A. P., 2012. Use of crash surrogates and exceedance statistics to estimate road safety. *Accident Analysis and Prevention, 45*, 230-240.

Tarko, A., Davis, G., Saunier, N., Sayed, T., & Washington, S., 2009. Surrogate Measures of Safety. *White Paper*.

Washington, S. P., Karlaftis, M. G., & Mannering, F. L., 2011. *Statistical and Econometric Methods for Transportation Data Analysis* (2nd ed.).

Wu, K.-F., & Jovanis, P. P., 2012. Crashes and crash-surrogate events: Exploratory modeling with naturalistic driving data. *Accident Analysis and Prevention, 45*, 507-516.

Wu, K.-F., & Jovanis, P. P., 2013. Defining and screening crash surrogate events using naturalistic driving data. *Accident Analysis and Prevention, 61*, 10-22.

Wu, K.-F., Aguero-Valverde, J., & Jovanis, P. P., 2014. Using naturalistic driving data to explore the association between traffic safety-related events and crash risk at driver level. *Accident Analysis and Prevention, 72*, 210-218.

Zegeer, C. V., Steward, J. R., Huang, H. H., Lagerwey, P. A., Feaganes, J., & Campbell, B. J., 2005. *Safety Effects of Marked versus Unmarked Crosswalks at Uncontrolled Locations.* Federal Highway Administration, U.S. Department of Transportation.