2015

Examining Multiple Approaches for the Transferability of Safety Performance Functions

Ahmed Tarek Ahmed Farid
University of Central Florida

EXAMINING MULTIPLE APPROACHES FOR THE TRANSFERABILITY OF SAFETY PERFORMANCE FUNCTIONS

by

AHMED FARID
B.Sc. American University in Dubai, 2012

A thesis submitted in partial fulfillment of the requirements
for the degree of Master of Science
in the Department of Civil, Environmental and Construction Engineering
in the College of Engineering and Computer Science
at the University of Central Florida
Orlando, Florida

Fall Term
2015

Major Professor: Mohamed Abdel-Aty

ABSTRACT

Safety performance functions (SPFs) are essential in road safety since they are used to predict crash frequencies. They are commonly applied for detecting hot spots in network screening and assessing whether road safety countermeasures are effective. In the Highway Safety Manual (HSM), SPFs are provided for several crash classifications for several types of roadway facilities. The SPFs of the HSM are developed using data from multiple states. In regions where jurisdiction specific SPFs are not available, it is custom to adopt nationwide SPFs for crash predictions then apply a calibration factor. Yet, the research is limited regarding the application of national SPFs for local jurisdictions. In this study, the topic of transferability is explored by examining rural multilane highway SPFs from Florida, Ohio, and California. That is for both divided segments and intersections. Traffic, road geometrics and crash data from the three states are collected to develop one-state, two-state and three-state SPFs. The SPFs are negative binomial models taking the form of those of the HSM. Evaluation of the transferability of models is undertaken by calculating a measure known as the transfer index. It is used to explain which SPFs may be transferred tolerably to other jurisdictions. According to the results, the transferability of rural divided segments' SPFs of Florida to California and vice versa is superior to that of Ohio's SPFs. For four-leg signalized intersections, neither state's models are transferable to any state. Also, the transfer index indicates improved transferability when using pooled data from multiple states. Furthermore, a modified version of the Empirical Bayes method that is responsible for segment specific adjustment factors is proposed as an alternative to the HSM calibration method. It is used to adjust crash frequencies predicted by the SPFs being transferred to the jurisdiction of interest. The modified method, proposed, outperforms the HSM calibration method as per the analysis results.

ACKNOWLEDGMENTS

I offer my sincere gratitude to my academic adviser and chair of the Civil, Environmental and Construction Engineering Department, Dr. Abdel-Aty. He was the first to direct me into the research area of the transferability of safety performance functions among states and provided great advice. In addition, I admit that I am indebted to declare my dear appreciation to my committee members, Dr. Jaeyoung Lee and Dr. Naveen Eluru. I thank Dr. Lee for his mentorship. Without him, I believe I could not have put in great effort. Also, whenever I needed help he was always available for consultation. He was attentive to me throughout the research. Special thanks are owed to Dr. Eluru for providing his exceptional advice, based on his expertise. Furthermore, the efforts of Ph.D. candidate Jung-Han Wang and Dr. Juneyoung Park are also appreciated for providing the analysis data. Finally, the financial sponsors of the National Cooperative Highway Research Program (NCHRP) project 17-62 ought to be greatly thanked. The effort of the project research team headed by Drs. John Ivan, Bhagwant Persaud, and Raghavan Srinivasan is highly appreciated.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER 1: INTRODUCTION

As per the Local and Rural Road Safety Program of the Federal Highway Administration (FHWA), in 2012, almost half of the fatal crashes in the US were on rural roads. Therefore, safety on rural roads is a crucial area in improving traffic safety. Safety performance functions (SPF) form a critical part of road safety improvement. Safety analysis need not necessarily be undertaken by applying SPFs. Other statistical methods such as the logistic regression approach for estimating crash risk (Pande and Abdel-Aty, 2009), matched case-control methods that require crash versus non-crash cases for identifying crash patterns (Jovanis and Gross, 2007), regression tree analysis for identifying the critical contributing factors to crashes and ordered probit models that are used to examine factors that contribute to severities of crashes (Abdel-Aty and Keller, 2005) are applied in traffic safety research. Driving simulator (Yan *et al.*, 2008) and crash worthiness (Huang *et al.*, 2011) studies have also contributed to the traffic safety literature. Crash worthiness is the potential for the victims of a crash not to be severely injured and recover as quickly as possible. In this context, the focus of the study is on SPFs. They are used to predict crash counts and identify hot spots on the road network typically for a forecast year. The SPFs may be implemented for prediction of the frequencies of crashes of any type or severity level. The crashes may be multi-vehicle (MV) crashes, head-on crashes, sideswipe crashes, rear-end crashes, left-turn crashes, angle crashes, pedestrian crashes, single-vehicle crashes (SV), bicycle crashes, animal crashes or any other type of crashes. The severity levels are: fatal (K), incapacitating injury (A), non-incapacitating injury (B), possible injury (C) and property damage only (O). The SPFs are developed by using the crash frequency, of the crash classification under study, as the dependent variable while modeling the traffic flow and roadway geometric features as independent variables. Ordinary linear regression models are problematic since crash counts are not continuous and cannot be negative (Miaou and Lum, 1993; Miaou, 1994; Kim *et al*., 2005; Garber and Wu, 2001). Generalized linear regression models (GLM) were applied lately (Sawalha and Sayed, 2006; Taylor *et al.*, 2002; Harnen *et al.*, 2004; Donnell and Mason, 2006). For instance, Ackaah and Salifu (2011) modeled crash frequencies on two lane roads in the Ashanti region of Ghana by means of the GLM. Poisson and Negative Binomial (NB) models are more appropriate mathematically since they account for the fact that crash frequencies are discrete. Yet, the mean and the variance of crash frequencies are typically unequal, a violation of the basic assumptions of the Poisson model. The variance, in most cases is greater than the mean, a condition known as overdispersion. Instead, in the current road safety literature, NB SPFs are fitted since the NB model accommodates the overdispersion (Miaou and Lum, 1993; Miaou, 1994; Harnen *et al.*, 2004; Lord *et al.*, 2005).

Negative Binomial SPFs are provided by the Highway Safety Manual (HSM) for several types of roadway facilities. Typically, the HSM's SPFs are for base conditions pertaining to specific roadway characteristics. For divided rural multilane highway segments, the base conditions are:

- 12 ft lanes
- 8 ft shoulders
- 30 ft medians
- No street lighting
- No automated speed enforcement

The SPFs for rural divided multilane highway segments in the HSM are developed based on crash data from California and Texas (Lord *et al.*, 2008). Specifically, both state's data are pooled and negative binomial SPFs are developed based on the combined data. In addition, four-leg signalized intersection SPFs are estimated based on data from Minnesota. Furthermore, development of separate SPFs for four-leg signalized intersections is undertaken by employing data from California. Based on the key findings and the research team's judgement, the recommended four-leg signalized intersection SPFs provided in the HSM are the estimated ones of Minnesota (Lord *et al.*, 2008). Therefore, as part of this study, the effect of pooling data from more than one state on the transferability of SPFs is examined. It should be noted that base conditions for four-leg signalized intersections are not defined in the current HSM as of now. It is common practice that if jurisdiction specific SPFs are unavailable, the national SPFs provided by the HSM are applied. Since roadway facilities in the jurisdiction of interest do not necessarily satisfy the HSM base conditions, crash modification factors (CMFs), also provided in the HSM, are used to adjust crash predictions accordingly but that is a topic beyond the scope of this study. The HSM's SPFs are then multiplied by calibration factors when applied to the jurisdiction of interest. The calibration factor is calculated as the ratio of the total observed crashes in all sites, whether segments or intersections, to the total predicted crashes in all sites. Thus, the calculated ratio is an aggregate factor. The multiplication of the predicted number of crashes by the calibration factor can be interpreted as adjusting the constant term of the SPF. The HSM's SPFs are applied to rural divided multilane highway segments and four-leg signalized intersections in several US states and in other nations. The observed crash frequencies are compared with those predicted by the HSM's SPFs to assess the SPFs' prediction accuracies. In addition, in some regions, jurisdiction specific SPFs are developed. They are compared with the HSM's SPFs.

This study contributes to the research on SPF transferability. Precisely, jurisdiction specific SPFs of Florida (FL), Ohio (OH) and California (CA) are estimated and applied to each state. The accuracies of the SPF predictions are assessed. Furthermore, data from multiple states are pooled and joint SPFs are developed. The joint SPFs are also applied to each state. That is to assess the effect of pooling the data from multiple states on transferability. In this study, the pooling of data from two states and three states for SPF estimation is considered. The SPFs are compared using a measure called the transfer index. In addition, a more disaggregate adjustment method is proposed to be used for correcting crash predictions of SPFs being transferred to the jurisdiction of interest. The HSM calibration method and the proposed one are compared in this study.

1.1 Objectives

- The aim of this study is to assess whether Florida, Ohio and California's SPFs, taking the form of those of the HSM, are transferable among each state for both rural divided multilane highway segments and four-leg signalized intersections. The SPFs are for total (KABCO), KABC, KAB, KA, SV and MV crashes. The assessment is conducted for average conditions, in which none of the variables is controlled. Furthermore, the assessment is undertaken for proposed modified versions of the HSM base conditions to examine the influence of controlling for variables on SPF transferability.
- Another objective is to investigate the impact of developing SPFs from pooled data of multiple states on the transferability of SPFs.
- The third objective is to assess the performance of the proposed modified Empirical Bayes Method relative to that of the HSM calibration factor method in terms of correcting the crash predictions.

In the following chapters, previous studies about SPF transferability, this study's data preparation, research methodology, analysis results interpretations, conclusions, key findings and suggestions for future work on SPF transferability assessment are discussed.

CHAPTER 2: LITERATURE REVIEW

In road safety literature, investigating the transferability of SPFs is a topic which is researched to a limited extent. Typically, the HSM's SPFs are applied to a specific jurisdiction and the calibration factors are calculated. Not in all of the previous studies jurisdiction specific SPFs are developed based on local data. The jurisdiction specific SPFs are compared with the HSM's SPFs that are multiplied by the calibration factors. This analysis approach has been the norm for studies of rural divided multilane highway segments, four-leg signalized intersections and other types of roadway facilities in Missouri, North Carolina, Oregon, Alabama, Regina, Saskatchewan, Canada and Toronto. This approach is also implemented abroad North America for the Messina-Catania region in Italy, Turin, Italy and Riyadh, Saudi Arabia. Another generic study was carried out to compare SPFs developed in the US with those of Sweden and New Zealand.

Sun *et al*. (2014) collected data of total crashes from the years 2009 to 2011 from the Transportation Management System (TMS) of the Missouri Department of Transportation (DOT) and applied the HSM's SPFs to rural divided multilane highway segments in Missouri. The HSM's NB model that takes the following form is used to predict the number of crashes for every rural divided multilane highway segment.

$$N_{SPF} = \exp[A + B\ln(AADT) + \ln(L)] \qquad (2.1)$$

In the SPF shown, L is the segment length, which is an exposure measure. Also, A and B are regression coefficients. The model with only the average annual daily traffic (AADT) and segment length is simple but convenient for identifying hot spots (Salifu, 2004). Since the segments, sampled for the Missouri study, do not necessarily conform to the HSM base conditions, the SPF being applied is multiplied by the HSM's CMFs. The calibration factor is determined to be 0.98 indicating that the HSM's SPF to a slight extent over-predicts frequencies of total crashes at rural divided multilane highway segments in Missouri. The study is also conducted for three and four-leg stop controlled intersections. According to the results, the calibration factors are less than 0.4 for both types of intersections. A possible explanation is that when the analysis is conducted, fewer than 100 crashes are observed for each intersection type causing more variability in the data. That is a direct violation to the HSM standards.

Srinivasan and Carter (2011) carried out the same type of analysis for rural divided multilane highway segments in North Carolina on data of total crashes that are collected between 2004 and 2008 from the Accident Analysis System of the North Carolina DOT. The HSM's SPF calibration factor is calculated to be 0.97. The analysis is also conducted for 19 four-leg signalized intersections and the crash data collected are from the same crash years. The resulting calibration factor is 0.49 for the intersections. Therefore, the HSM's SPF for four-leg signalized intersections over-predicts total crashes by 51%.

Xie *et al*. (2011) predicted frequencies of total crashes for rural divided multilane highway segments and other types of roadway facilities in Oregon by processing crash data from the years 2004 to 2006. According to the results of the analysis, the calibration factor for the divided segments is 0.78 after taking into consideration the HSM's CMFs. That indicates that the total crashes in Oregon are to a considerable extent over-predicted by the nationwide HSM SPF. Variations in crash reporting thresholds are potential explanations. Since part of the data processed for development of rural multilane highway segments of the HSM are from California and Texas, these states' reporting thresholds are different from that of Oregon. The crash reporting threshold of Oregon is $1,500 worth of property damage (Xie *et al*., 2011). That is twice that of California, which is $750 (Xie *et al*., 2011), and 50% greater than that of Texas, which is $1,000 as per the Driver Safety and Laws policy of the Texas Department of Transportation.

In Alabama, Mehta and Lou (2013), not only applied the HSM's SPFs to the local jurisdiction but also compared the HSM's SPFs' performances with local SPFs for rural divided multilane highway segments. In the study, records of total crashes from 2006 to 2009 are used to develop local SPFs. The research team estimated a series of SPFs having different functional forms. In addition, the HSM's SPFs are applied and multiplied by calibration factors after taking into consideration the CMFs also provided in the HSM. The local SPFs, estimated, are compared with those of the HSM that are multiplied by calibration factors when applied to a validation dataset of 2000 homogeneous segments. The mean absolute deviation (MAD), mean predicted bias (MPB), mean squared predicted error (MSPE), and Akaike's information criterion (AIC) are used to compare the SPFs. The superior SPF is a local one similar to that of the HSM but with supplementary variables including the lane width, posted speed limit and a dummy variable that represents the crash year. The superior model also includes a coefficient for the segment length as opposed to the HSM's SPF where the segment length is not associated with any coefficient.

In another study by Young and Park (2012), HSM SPFs are calibrated for signalized and unsignalized intersections in Regina, Saskatchewan, Canada. The intersections include 143 signalized intersections, 123 three-leg unsignalized intersections and 121 four-leg unsignalized intersections. The crash data are collected from the province's government insurance records. Crash years are from 2005 to 2009. The geometric data are collected in the form of shapefiles accessible by means of ArcGIS software. From the crash data, 70% are used for estimation of local SPFs while the remaining 30% are used for validation. The models are for total, KABC and property damage only crashes. The local SPFs take the following form:

$$N_{SPF} = \exp(\beta_0) \times (AADT_{major}/1000)^{\beta_1} \times (AADT_{minor}/1000)^{\beta_2} \tag{2.2}$$

The terms $AADT_{major}$ and $AADT_{minor}$ are the major and minor road entering AADTs respectively. The local SPFs' prediction accuracies are compared with those of the HSM's SPFs and of the HSM's SPFs multiplied by the calibration factors. In terms of MAD, MSPE and Freeman-Tukey's R^2 measures, the local SPFs outperform those of the HSM.

Other than the study in Regina, Saskatchewan, Canada, Hadayeghi *et al*. (2006) investigated the temporal transferability of SPFs of Toronto. That is, the authors studied whether the Toronto SPFs developed in 1996 are applicable for an extended period of time towards the year 2001 at a macroscopic level for every zone under study. The models are NB SPFs developed for total crashes, total crashes during the morning peak period, fatal and injury (FI) crashes, and FI crashes for the morning peak period. The variables used are: the natural logarithm of the vehicle-kilometers traveled, sum of the road lengths in each zone in kilometers, number of households in thousands, number of employments in thousands, speed, volume to capacity ratio and density of intersections. An approach known as the Bayesian updating method is used to modify the 1996 SPFs. The method is first proposed and applied by Atherton and Ben-Akiva (1976). It is used to update the variable coefficients of the 1996 Toronto models based on the 1996 variable coefficients, 1996 variable standard deviations, coefficients developed from data sampled in 2001 and standard deviations of the coefficients developed from the data sampled in 2001. Also, the authors applied a calibration factor to facilitate the transferability of the unadjusted 1996 models to 2001 as a separate part of the study. The factor is calculated as the ratio of the sum of the observed crashes in 2001 to the sum of the crashes in 1996 that is predicted by the SPF developed from the 2001 data. The 1996 SPFs are multiplied by the calibration factors and the performances of SPF predictions are compared with those of the 1996 SPFs updated by the Bayesian method. The measures used for comparison are MPB, MAD and mean square error (MSE). According to the results, the calibration factor method outperforms the Bayesian updating method. In addition, the transfer index (Hadayeghi *et al.*, 2006; Sikder *et al.*, 2014) and the nested log-likelihood measures are applied to compare the 1996 and 2001 models' log-likelihoods. According to the results of the nested log-likelihood measure, the 1996 models are not transferable to 2001. Yet, the transfer index results indicate that the 1996 SPF for FI crashes and FI crashes during the morning peak account for more than 50% of the variability in severe crash patterns in 2001. That is most likely because the vehicle-kilometers traveled and the demographic factors for both 1996 and 2001 are not significantly different as per results of the t-test conducted in the study.

In another study conducted in Toronto, Canada, Persaud *et al*. (2002) estimated SPFs for urban three-leg and four-leg intersections. The intersections are signalized and unsignalized intersections. The SPFs are NB models developed for total, injury and property damage only crashes using crash data from the years 1990 to 1995. Several functional forms of the SPFs are applied. One includes the natural logarithm of the major road AADT, the natural logarithm of the minor road AADT and the major road AADT without any transformation as variables. Another is similar in form but instead of having the third variable as the major road AADT, the minor road AADT is included without being transformed. The third functional form is similar to that of the two SPFs, described, but without the third variable, which is either the major or minor road AADT. Local SPFs are developed for the intersections in Toronto, Vancouver, Canada and

California separately. The ones of Vancouver and California are multiplied by calibration factors to be applied for Toronto's conditions. The calibration factor is calculated in the same manner as that of the HSM as the ratio of the sum of the observed crash frequencies in Toronto to that of the ones predicted by the transferred model, whether that of Vancouver or California. The calibration factors for the SPFs of both Vancouver and California range from 1.2 to 1.8 indicating that the SPFs markedly under-predict crashes in Toronto. However, the root mean square errors (RMSE) of the SPFs of the three regions are to an extent similar. The research team used cumulative residual plots (Hauer and Bamfo, 1997) for the entering AADTs and concluded that the fits of the SPFs multiplied by the calibration factors are satisfactory for the range of entering AADTs.

The application and calibration of SPFs of the HSM is also undertaken globally. Cafiso *et al*. (2012) undertook a similar study for the A18 divided multilane highway in the Italian region, Messina-Catania by analyzing KAB crash records collected from the years, 2005 to 2008. The research team calculated calibration factors for each year separately while taking into consideration the CMFs. The range of calibration factors is from 1.14 to 1.43. The average of the calibration factors is 1.26 indicating that the HSM's SPF under-predicts severe crash frequencies for Messina-Catania's divided segments by 26%. In addition, two local SPFs are developed. One is a simple model which only includes the natural logarithm of the AADT as a variable and the natural logarithm of the segment length without a coefficient. The other is a more complicated multivariable one with not only the transformed AADT and segment length but also horizontal curvature and gradient variables. Both local SPFs are compared in terms of the adjusted coefficient of determination and Pearson's chi-squared statistic. The latter model performed better. Both the local multivariable model and the HSM's SPF multiplied by the calibration factor are used in network screening and compared. The RMSE is used as a relative measure of the difference among predictions of both SPFs. According to the results, the predictions of both SPFs are not considerably different.

Other than the study in Messina-Catania, Sacchi *et al*. (2012) conducted a study in Turin, Italy to investigate the transferability of the HSM's SPFs and CMFs for rural two-lane roads. Traffic volume, road geometrics and crash data are collected from the Italian National Institute of Statistics. The crashes are FI crashes of which records are collected from the years 2005 to 2008. The data are refined to exclude segments not satisfying the HSM base conditions for rural two lane roads. Both the HSM's provided SPF and a locally developed one are applied and the resulting calibration factor is calculated to be 0.44. That is an implication that the HSM model over-predicts FI crashes by 56%. Yet, for low AADTs the predictions of both SPFs are similar.

In another study conducted abroad the US by Al Kaaf and Abdel-Aty (2015), crash and geometric data are collected from urban divided roads in Riyadh, Saudi Arabia. The HSM models, used for predicting fatal and injury (FI) crashes are applied. The CMFs, provided by the HSM, are also applied. In the results of the analysis, the calibration factor is calculated as 0.31 implying that the HSM, to a great extent, over-predicts crash frequencies in Riyadh, Saudi Arabia's urban divided roads. Also, several local SPFs having different functional forms are developed and the superior one is a model with the variables including: the natural logarithm of the AADT, segment length with an associated coefficient, posted speed limit and driveway density. Furthermore, the research team developed CMFs using local data of Riyadh, Saudi Arabia and investigated the performances of the HSM's CMFs relative to those of the CMFs developed. The local CMFs produced improved results.

Turner *et al*. (2007) compared total crashes' SPFs for roundabouts in New Zealand, Sweden and the US. The variable included in the SPFs is the natural logarithm of the total entering AADT. Yet, the crash years based on which each nation's SPF is developed are different for the other regions. Generally, the predictions of all three nation's SPFs are similar for AADTs less than 5000 vpd. For AADTs between 20,000 vpd to 25,000 vpd, Sweden and the US's SPF predictions are similar as well. For all other AADTs the SPF predictions are considerably different. Variations may be explained by different reporting thresholds, weather conditions, posted speed limits and how intersection influence areas are defined in every nation. The authors also developed SPFs for total crashes at rural two-lane roads in New Zealand, Minnesota, North Carolina, Ohio and Washington. Calibration factors are calculated for the New Zealand model to be applied to each state in the US mentioned. The factors are also calculated for the US states to be applied amongst each other and to New Zealand. The calibration factors of New Zealand range from 1.3 to 3.85. Thus, the New Zealand model under-predicts crashes in the US by a great extent. The research team also implemented cumulative residual plots (Hauer and Bamfo, 1997) using the AADT as the variable to compare the US states' SPFs multiplied by calibration factors with that of New Zealand under the condition that the jurisdiction where the SPFs are applied is New Zealand. From the plots, the New Zealand, Minnesota and Washington cumulative residuals are similar throughout the range of AADTs.

In general, calibration of the predicted crash frequencies obtained from the HSM's SPFs for specific jurisdictions is a subject of investigation by a considerable number of researchers in road safety. Yet, there is a growing need for research on the topic of transferability of SPFs. Furthermore, in none of the previously stated studies, except for that aimed at formulating the national SPFs of the HSM, the authors proposed and conducted a methodical analysis to investigate the advantage of pooling data from several states when developing SPFs. The rural divided and undivided highway segment SPFs of the HSM are developed from pooled data mainly to ensure sufficient sample sizes (Lord *et al*., 2008). In addition, in the previously stated studies, the HSM calibration factors are simply applied to adjust crash predictions. Setbacks of the previous studies are addressed. First, the benefit of estimating SPFs from combined data from several states, Florida, Ohio and California, is explored and quantified. The SPFs, based on the pooled data, are modified to include dummy variables representing the states. That also applies for the over-dispersion parameters of the SPFs. Second, a transfer index is applied to measure the performance of SPFs that are being transferred to jurisdictions of interest. Third, a method is proposed and studied for adjusting predicted crash frequencies at every segment or intersection. Besides, in the HSM calibration factor method, the calibration factors are based on an aggregate measure.

CHAPTER 3: DATA PREPARATION FOR RURAL DIVIDED MULTILANE HIGHWAY SEGMENTS

The Florida crash data are collected from the Crash Analysis Reporting System (CARS) of the Florida Department of Transportation (FDOT). The road geometric data of Florida are collected from the Roadway Characteristics Inventory (RCI) of the FDOT. The Ohio and California data, including traffic data, crash records and road characteristics, are collected from the Highway Safety Information System (HSIS) of the Federal Highway Administration (FHWA). The HSIS database not only contains data of Ohio and California but also of Washington, Minnesota, Illinois, Michigan, North Carolina and Maine. In the following subsections, the data preparation step for rural divided multilane highway segments is explained. The data preparation step for the four-leg signalized intersections is explained in the following chapter. For both segments and intersections, pedestrian, bicycle and animal crashes are excluded from all three states' data.

3.1 Rural Divided Multilane Highway Segments Data Preparation

The Florida, Ohio and California data are prepared once for average conditions, which are those not conforming to the HSM base conditions and proposed modified base conditions. The modified base conditions are:

- Lane width ≥ 12 ft
- Paved shoulder of which width ≥ 8 ft
- Median width ≥ 30 ft
- No street lighting
- No automated speed enforcement

The reason the modified HSM base conditions are applied is that there is an inadequate sample of segments that satisfy the base conditions in each state for SPF development. In the following subsections, the data preparation is described for average and modified base conditions of rural divided multilane highway segments.

3.1.1 Rural Divided Multilane Highway Segments Data Preparation – Average Conditions

The Florida data collected are composed of records of crashes that occurred from 2009 to 2011. The number of segments of which geometric characteristics are available in the Florida data is 1,320. The segments comprise 835.86 mi. The Ohio data also encompass 2009 to 2011 crash records. Ohio's segments data include geometric features of 1,261 segments having an aggregate length of 665.11 mi. Unlike Florida and Ohio, the California data comprise crash records from the years 2009 to 2010. The road geometrics data of California are collected from 1,349 homogenous segments comprising 709.83 mi. In accordance with the HSM, the minimum segment length, for which data are collected, is not less than 0.1 mi in any of the three states. The crashes are classified into the categories: KABCO, KABC, KAB, KA, single vehicle (SV) and multi-vehicle (MV) crashes. Descriptive statistics of the Florida, Ohio and California data are shown in Tables 3.1 through 3.3.

Table 3.1: Florida's Average Conditions Descriptive Statistics for Rural Divided Multilane Highway Segments

Number of Segments = 1,320	Number of Crashes	Mean	Standard Deviation
Segment Length (mi)	-	0.633	0.993
AADT (vpd)	-	25,710.482	12,001.344
Lane Width (ft)	-	11.845	0.489
Shoulder Width (ft)	-	4.225	2.277
Median Width (ft)	-	28.258	18.114
KABCO	10,028	7.597	15.001
KABC	4,815	3.648	6.486
KAB	2,399	1.817	3.094
KA	753	0.57	1.146
SV	1,929	1.461	2.364
MV	8,099	6.136	14.055

Table 3.2: Ohio's Average Conditions Descriptive Statistics for Rural Divided Multilane Highway Segments

Number of Segments = 1,261	Number of Crashes	Mean	Standard Deviation
Segment Length (mi)	-	0.527	0.582
AADT (vpd)	-	9,896.954	5,600.405
Lane Width (ft)	-	11.733	0.484
Shoulder Width (ft)	-	6.452	2.504
Median Width (ft)	-	43.41	21.616
KABCO	2,541	2.015	4.028
KABC	799	0.634	1.507
KAB	580	0.46	1.069
KA	145	0.115	0.382
SV	1,362	1.08	2.094
MV	1,179	0.935	2.605

Table 3.3: California's Average Conditions Descriptive Statistics for Rural Divided Multilane Highway Segments

Number of Segments = 1,349	Number of Crashes	Mean	Standard Deviation
Segment Length (mi)	-	0.526	0.568
AADT (vpd)	-	19,018.744	14,370.401
Lane Width (ft)	-	10.088	4.426
Shoulder Width (ft)	-	7.681	2.377
Median Width (ft)	-	36.787	31.262
KABCO	5,120	3.795	6.805
KABC	1,997	1.48	2.734
KAB	1,014	0.752	1.392
KA	283	0.21	0.54
SV	2,170	1.609	3.273
MV	2,950	2.187	4.532

As shown in the tables, the lane width means and standard deviations in all three states are not similar. California’s lane widths’ mean and standard deviation are different from those of the other states’ lane widths. That is a factor that may deter the transferability of SPFs of Florida and Ohio to California. The shoulder widths’ mean is also different for each state even though the standard deviations are similar and low. That indicates that there is a low degree of variability in shoulder widths in each state’s data. Specifically, the state of Florida has the narrowest mean shoulder width, followed by Ohio followed by California. That is a factor than may inhibit SPF transferability among the three states. On the other hand, the median widths’ means are different for every state. Also, the standard deviation of the median widths of California are high. This indicates that there is a high degree of variability in California’s median widths relative to those of Florida and Ohio. This inhibits transferability of SPFs of Florida and Ohio to California. The

AADTs' means and standard deviations of all states' data are different. Therefore, it is expected that each of the three states experiences different crash frequencies which are proportional to the AADT because it is an exposure measure. It should be noted that there are discrepancies in the data. Less than 2% of the crashes in Florida are missing severity levels. Also, less than 0.75% of the crashes in California are without severity level codes. These types of crashes are considered for modeling predictions of KABCO crashes, SV and MV crashes only. Finally, all state's crashes are normalized by the hundred million vehicle miles traveled (VMT) per year as shown in Table 3.4. The normalized crashes are termed crash rates.

Table 3.4: Florida, Ohio and California's Average Conditions Rural Divided Segment Crash Rates per Hundred Million Vehicle Miles Traveled per Year

Crash Classification	Florida	Ohio	California
KABCO	47.70	32.75	53.93
KABC	22.91	10.3	21.03
KAB	11.41	7.48	10.68
KA	3.58	1.87	2.98
SV	9.18	17.55	22.86
MV	38.53	15.2	31.07

As, shown in Table 3.4, the least KABCO, KABC, KAB, KA and MV crashes per hundred million VMT are experienced by Ohio. The crash rates of both Florida and California are reasonably higher than those of Ohio. It is crucial to note that KABCO, KABC, KAB and KA crash rates of Florida and California are similar. Therefore, Florida's SPFs for these types of crashes are expected to be transferable to California and vice versa. However, California experiences the most SV crashes per hundred million VMT per year followed by Ohio which is followed by Florida. The data of the three states are pooled into different combinations to be later used for SPF development. That is, SPFs are developed from pooled data of two states and three states. The pooled data descriptive statistics are shown in Tables 3.5 through 3.8.

Table 3.5: Florida and Ohio's Joint Data Average Conditions Descriptive Statistics for Rural Divided Multilane Highway Segments

Number of Segments = 2,581	**Number of Crashes**	**Mean**	**Standard Deviation**
Segment Length (mi)	-	0.582	0.82
AADT (vpd)	-	17,984.461	12,306.954
Lane Width (ft)	-	11.79	0.49
Shoulder Width (ft)	-	5.313	2.637
Median Width (ft)	-	35.661	21.291
KABCO	12,569	4.87	11.435
KABC	5,614	2.175	4.989
KAB	2,979	1.154	2.432
KA	898	0.348	0.891
SV	3,291	1.275	2.244
MV	9,278	3.595	10.539

Table 3.6: Florida and California's Joint Data Average Conditions Descriptive Statistics for Rural Divided Multilane Highway Segments

Number of Segments = 2,669	**Number of Crashes**	**Mean**	**Standard Deviation**
Segment Length (mi)	-	0.579	0.809
AADT (vpd)	-	22,328.258	13,665.354
Lane Width (ft)	-	10.957	3.284
Shoulder Width (ft)	-	5.972	2.899
Median Width (ft)	-	32.569	25.965
KABCO	15,148	5.676	11.759
KABC	6,812	2.552	5.075
KAB	3,413	1.279	2.449
KA	1,036	0.388	0.911
SV	4,099	1.536	2.86
MV	11,049	4.14	10.58

Table 3.7: Ohio and California's Joint Data Average Conditions Descriptive Statistics for Rural Divided Multilane Highway Segments

Number of Segments = 2,610	Number of Crashes	Mean	Standard Deviation
Segment Length (mi)	-	0.527	0.575
AADT (vpd)	-	14,611.626	11,942.810
Lane Width (ft)	-	10.883	3.303
Shoulder Width (ft)	-	7.087	2.515
Median Width (ft)	-	39.987	27.231
KABCO	7,661	2.935	5.706
KABC	2,796	1.071	2.267
KAB	1,594	0.611	1.255
KA	428	0.164	0.473
SV	3,532	1.353	2.779
MV	4,129	1.582	3.779

Table 3.8: Florida, Ohio and California's Joint Data Average Conditions Descriptive Statistics for Rural Divided Multilane Highway Segments

Number of Segments = 3,930	Number of Crashes	Mean	Standard Deviation
Segment Length (mi)	-	0.563	0.744
AADT (vpd)	-	18,339.486	13,059.478
Lane Width (ft)	-	11.206	2.744
Shoulder Width (ft)	-	6.126	2.787
Median Width (ft)	-	36.047	25.165
KABCO	17,689	4.501	10.1
KABC	7,611	1.937	4.361
KAB	3,993	1.016	2.141
KA	1,181	0.301	0.791
SV	5,461	1.39	2.647
MV	12,228	3.111	8.968

3.1.2 Rural Divided Multilane Highway Segments Data Preparation – Modified Base Conditions

The prepared rural divided multilane highway segment data of all three states are refined to reflect the proposed modified HSM base conditions. Specifically, the number of segments that satisfy the modified base conditions in Florida, Ohio and California are 57, 432 and 572 respectively. The descriptive statistics of the three states' modified base conditions data are shown in Tables 3.9 through 3.11.

Table 3.9: Florida's Modified Base Conditions Descriptive Statistics for Rural Divided Multilane Highway Segments

Number of Segments = 57	**Number of Crashes**	**Mean**	**Standard Deviation**
Segment Length (mi)	-	0.509	0.64
AADT (vpd)	-	30,275.667	12,535.496
Lane Width (ft)	-	12.018	0.093
Shoulder Width (ft)	-	9.684	2.354
Median Width (ft)	-	44.509	17.878
KABCO	290	5.088	6.911
KABC	152	2.667	4.142
KAB	93	1.632	2.907
KA	38	0.667	1.418
SV	95	1.667	2.452
MV	195	3.421	4.866

Table 3.10: Ohio's Modified Base Conditions Descriptive Statistics for Rural Divided Multilane Highway Segments

Number of Segments = 432	**Number of Crashes**	**Mean**	**Standard Deviation**
Segment Length (mi)	-	0.588	0.651
AADT (vpd)	-	11,188.016	4,556.527
Lane Width (ft)	-	12	0
Shoulder Width (ft)	-	8	0
Median Width (ft)	-	47.87	16.605
KABCO	917	2.123	3.407
KABC	262	0.606	1.252
KAB	193	0.447	0.967
KA	48	0.111	0.336
SV	554	1.282	2.204
MV	363	0.84	1.693

Table 3.11: California's Modified Base Conditions Descriptive Statistics for Rural Divided Multilane Highway Segments

Number of Segments = 572	Number of Crashes	Mean	Standard Deviation
Segment Length (mi)	-	0.569	0.584
AADT (vpd)	-	19,488.747	13,447.115
Lane Width (ft)	-	12.043	1.004
Shoulder Width (ft)	-	9.047	1.002
Median Width (ft)	-	65.11	22.078
KABCO	2,004	3.503	6.845
KABC	758	1.325	2.502
KAB	398	0.696	1.304
KA	116	0.203	0.53
SV	900	1.573	2.809
MV	1,104	1.93	4.906

As shown in Tables 3.9 through 3.11, the means of the lane and shoulder widths of Florida's segments conforming to the modified base conditions are similar to those of Ohio and California's lane widths. The standard deviations are also similar except for that of California's lane widths. The variability in California's lane widths is to a limited extent greater than those of Florida and Ohio's lane widths. The shoulder widths' means ranked from widest to narrowest are those of Florida, California and Ohio. Also there is a great degree of variability in Florida's shoulder widths relative to California's as indicated by the standard deviations. Ohio's shoulder widths are all the same since the standard deviation is null. Therefore, it is anticipated that the differences in lane widths among divided segments in all three states will not impede the transferability of SPFs as will the differences in shoulder widths. On the other hand, the means of the median widths of Florida and Ohio's segments conforming to the modified base conditions are considerably lower than that of California. Yet, the standard deviations of the median widths in all three state's segments are high. This is an indication of a large degree of variability in median widths in each state. That is a factor that may inhibit transferability of modified base conditions SPFs for divided segments among the three states. It should be noted that less than

2% of the crashes that occurred at segments conforming to the modified base conditions in Florida and less than 1% of those in California are crashes of which the severities are unknown. They cannot be used for modeling of KABC, KAB and KA crashes. The crash rates of all three states are shown in Table 3.12.

Table 3.12: Florida, Ohio and California's Modified Base Conditions Rural Divided Segment Crash Rates Per Hundred Million Vehicle Miles Traveled Per Year

Crash Classification	Florida	Ohio	California
KABCO	28.46	29.17	42.63
KABC	14.92	8.33	16.12
KAB	9.13	6.14	8.47
KA	3.73	1.53	2.47
SV	9.32	17.62	19.15
MV	19.14	11.55	23.49

As shown in Table 3.12, California's crash rates are considerably larger than both Florida and Ohio. Yet, California's crash rates are similar to Florida's for KABC and KAB crashes. Also, Ohio experiences the least KABC, KAB, KA and MV crashes per hundred million VMT. It is critical to note that after the lane width, shoulder width, median width, lighting conditions and speed enforcement variables are controlled, the KABCO crash rates of Florida and Ohio become similar. It is expected that the KABCO modified base conditions SPF for segments in Florida is transferable to Ohio and vice versa as opposed to the case of the average conditions. As is the case of the segments' average conditions, each state's data of the segments' modified base conditions are pooled with the other states' data in different combinations. However, only Ohio and California's segments' modified base conditions data can be pooled since Florida's data sample size is low relative to that of the data of the other states. The descriptive statistics of the pooled Ohio and California's modified base conditions segments data are shown in Table 3.13.

Table 3.13: Ohio and California's Modified Base Conditions Descriptive Statistics for Rural Divided Multilane Highway Segments

Number of Segments = 1,004	Number of Crashes	Mean	Standard Deviation
Segment Length (mi)	-	0.577	0.613
AADT (vpd)	-	15,917.118	11,347.743
Lane Width (ft)	-	12.024	0.758
Shoulder Width (ft)	-	8.597	0.917
Median Width (ft)	-	57.692	21.655
KABCO	2,921	2.909	5.669
KABC	1,020	1.016	2.089
KAB	591	0.589	1.177
KA	164	0.163	0.459
SV	1,454	1.448	2.569
MV	1,467	1.461	3.902

CHAPTER 4: DATA PREPARATION FOR RURAL FOUR-LEG SIGNALIZED INTERSECTIONS

The data sources for all three states for the four-leg signalized intersections are the same as those of the segments. As is the case of the segments, the four-leg signalized intersections are processed for average conditions and base conditions. However, base conditions are not defined for four-leg signalized intersections in the current HSM. The proposed base conditions for four-leg signalized intersections are:

- Skew angle between 0^{o} and 5^{o}
- Street lighting presence
- No red-light-running cameras

An additional condition specified by the HSM for stop controlled intersections is that the intersections should have no turning lanes. This condition is neglected in the proposed base conditions because there is an inadequate sample of signalized intersections in each state with no turning lanes. Also, in accordance with the HSM standards, any crashes within a 250 ft radius of intersections are considered intersection or intersection related crashes. Descriptions of the data preparation for average and proposed base conditions for four-leg signalized intersections are provided in the following subsections. Similar to the case of the segments, data corresponding to crash records from 2009 to 2011 are collected for Florida and Ohio. In the case of California, records of crashes that occurred in 2009 and 2010 are collected for analysis. The 2009 and 2010 crash data of Florida and Ohio are to be processed for SPF development while the 2011 crash data of both states will be used for application of the proposed modified Empirical Bayes method discussed in the research methodology chapter. For California, the 2009 crash data are used for estimating SPFs and the modified Empirical Bayes method will be applied to the 2010 data. The crashes are classified the same way for analysis purposes as they are for the case of the segments.

4.1 Rural Four-Leg Signalized Intersections Data Preparation – Average Conditions

There are 131 rural four-leg signalized intersections sampled from Florida. The software ArcMap is used to geocode crash locations within the 250 ft buffers of the four-leg signalized intersections. The geocoding is performed to aid in matching the Florida crash data from the CARS database with the road log data from the RCI database. In Ohio and California, there are 122 and 34 four-leg signalized intersections sampled, respectively. The descriptive statistics for each state's average conditions four-leg signalized intersection data, excluding crash data used for the modified Empirical Bayes method, are shown in Tables 4.1 through 4.3. For the California four-leg signalized intersection average conditions data, no data is available regarding the skew angle.

Table 4.1: Florida's Average Conditions Descriptive Statistics for Rural Four-Leg Signalized Intersections

Number of Intersections =131	**Number of Crashes**	**Mean**	**Standard Deviation**	
Major Road AADT	-	11854.97	5874.483	
Minor Road AADT	-	4737.752	3143.863	
Total AADT	-	16592.72	7660.074	
Skew Angle	-	13.996	21.933	
Presence of RTLs	-	0.672	0.471	
Presence of LTLs	-	0.908	0.29	
Lighting	-	0.427	0.497	
KABCO	134	1.023	1.591	
KABC	71	0.542	0.987	
KAB	34	0.26	0.549	
KA	10	0.076	0.267	
SV	7	0.053	0.226	
MV	127	0.969	1.549	
Frequencies for Categorical Variables				
	Frequency Present	**Percent Present**	**Frequency Absent**	**Percent Absent**
Presence of RTLs	88	67.18	43	32.82
Presence of LTLs	119	90.84	12	9.16
Lighting	56	42.75	75	57.25

Table 4.2: Ohio's Average Conditions Descriptive Statistics for Rural Four-Leg Signalized Intersections

Number of Intersections =122	**Number of Crashes**	**Mean**	**Standard Deviation**	
Major Road AADT	-	9927.074	7117.279	
Minor Road AADT	-	2622.549	2542.288	
Total AADT	-	12549.62	7905.492	
Skew Angle	-	9.91	13.384	
Presence of RTLs	-	0.164	0.372	
Presence of LTLs	-	0.525	0.501	
Lighting	-	0.73	0.446	
KABCO	544	4.459	4.128	
KABC	172	1.41	2.394	
KAB	102	0.836	1.597	
KA	18	0.148	0.492	
SV	28	0.23	0.542	
MV	516	4.23	4.057	
Frequencies for Categorical Variables				
	Frequency Present	**Percent Present**	**Frequency Absent**	**Percent Absent**
Presence of RTLs	20	16.39	102	83.61
Presence of LTLs	64	52.46	58	47.54
Lighting	89	72.95	33	27.05

Table 4.3: California's Average Conditions Descriptive Statistics for Rural Four-Leg Signalized Intersections

Number of Intersections =34	**Number of Crashes**	**Mean**	**Standard Deviation**	
Major Road AADT	-	22504.06	11110.56	
Minor Road AADT	-	4026.912	4438.882	
Total AADT	-	26530.97	12143.41	
Presence of RTLs	-	0.706	0.462	
Presence of LTLs	-	0.97	0.174	
Lighting	-	1	0	
KABCO	106	3.118	2.783	
KABC	47	1.382	1.206	
KAB	19	0.559	0.746	
KA	4	0.118	0.327	
SV	13	0.382	0.779	
MV	93	2.735	2.514	
Frequencies for Categorical Variables				
	Frequency Present	**Percent Present**	**Frequency Absent**	**Percent Absent**
Presence of RTLs	24	70.59	10	29.41
Presence of LTLs	32	96.97	1	3.03
Lighting	34	100.00	0	0.00

The mean of the major road AADTs of four-leg signalized intersections in California is double those of Florida and Ohio's major road AADTs. Yet the standard deviations of the major road AADTs in intersections in all three states are high especially in California. The means and standard deviations of minor road AADTs of intersections in Florida and California are similar even though the standard deviations are considerably large. Ohio's intersections' minor road AADTs' mean and standard deviation are lower than those of Florida and California's minor road AADTs. The high standard deviations of major and minor road AADTs indicate high degrees of variability in entering traffic volumes and hence crashes. The AADTs from major and minor roads are factors that impede SPF transferability. The skew angle standard deviations of both the Florida and Ohio intersections are reasonably high indicating a high degree of variability. This may also inhibit transferability of SPFs among Florida and Ohio. In addition, the proportions of right turn lanes, left turn lanes and lighting are different for each state's four-leg signalized intersection data. These are other factors that may deter SPF transferability among the three states. The intersection crash rates are calculated as per hundred million VMT and are shown in Table 4.4.

Table 4.4: Florida, Ohio and California Average Conditions Rural Four-Leg Signalized Intersection Crash Rates Per Hundred Million Vehicle Miles Traveled Per Year

Crash Classification	Florida	Ohio	California
KABCO	5.63	32.45	16.10
KABC	2.98	10.26	7.14
KAB	1.43	6.08	2.89
KA	0.42	1.07	0.61
SV	0.29	1.67	1.97
MV	5.34	30.78	14.12

In general, Florida experiences the least crashes at four-leg signalized intersections per hundred million VMT per year followed by California, followed by Ohio except for SV crashes as shown in Table 4.4. The SV crash rates of California are higher than those of Ohio. Since there are no similarities in crash rates among states it is not expected that the SPFs of each state are transferable to either states. Furthermore, as is the case of segments, average conditions four-leg signalized intersection data of Florida and Ohio are pooled for analysis purposes. California's intersection data cannot be pooled with those of any other state since its sample size is low relative to the other states leading to biased results. The descriptive statistics of the pooled Florida and Ohio intersection data are shown in Table 4.5.

Table 4.5: Florida and Ohio's Average Conditions Descriptive Statistics for Rural Four-Leg Signalized Intersections

Number of Intersections = 253	Number of Crashes	Mean	Standard Deviation	
Major Road AADT	-	10925.31	6561.783	
Minor Road AADT	-	3717.773	3053.48	
Total AADT	-	14643.08	8023.453	
Skew Angle	-	12.018	18.379	
Presence of RTLs	-	0.427	0.496	
Presence of LTLs	-	0.723	0.448	
Lighting	-	0.573	0.496	
KABCO	678	2.68	3.528	
KABC	243	0.96	1.855	
KAB	136	0.538	1.21	
KA	28	0.111	0.393	
SV	35	0.138	0.419	
MV	643	2.542	3.436	
Frequencies for Categorical Variables				
	Frequency Present	**Percent Present**	**Frequency Absent**	**Percent Absent**
Presence of RTLs	108	42.69	145	57.31
Presence of LTLs	183	72.33	70	27.67
Lighting	145	57.31	108	42.69

4.2 Rural Four-Leg Signalized Intersections Data Preparation – Proposed Base Conditions

Florida and Ohio's average conditions four-leg signalized intersections are subset to include only intersections with street lighting, no automated speed enforcement and skew angles between 0^o and 5^o. Also, California's data of the four-leg signalized intersections conforming to the base conditions, proposed, are not prepared since no information is known about the skew angles. There are 39 and 63 four-leg signalized intersections that satisfy the proposed base conditions in Florida and Ohio, respectively. The descriptive statistics of both states' data are shown in Tables 4.6 through Table 4.7.

Table 4.6: Florida's Base Conditions Descriptive Statistics for Rural Four-Leg Signalized Intersections

Number of Intersections = 39	Number of Crashes	Mean	Standard Deviation	
Major Road AADT	-	11,372.731	6,176.434	
Minor Road AADT	-	4,027.795	2,929.942	
Total AADT	-	15,400.526	7,992.558	
Skew Angle	-	0	0	
Presence of RTLs	-	0.462	0.505	
Presence of LTLs	-	0.821	0.389	
Lighting	-	1	0	
KABCO	35	0.897	1.553	
KABC	15	0.385	0.633	
KAB	6	0.154	0.366	
KA	1	0.026	0.16	
SV	1	0.026	0.16	
MV	34	0.872	1.508	
Frequencies for Categorical Variables				
	Frequency Present	**Percent Present**	**Frequency Absent**	**Percent Absent**
Presence of RTLs	18	46.15	21	53.85
Presence of LTLs	32	82.05	7	17.95
Lighting	39	100	0	0.00

Table 4.7: Ohio's Base Conditions Descriptive Statistics for Rural Four-Leg Signalized Intersections

Number of Intersections = 63	Number of Crashes	Mean	Standard Deviation	
Major Road AADT	-	7,584.444	4,921.543	
Minor Road AADT	-	2,678.937	2,643.804	
Total AADT	-	10,263.38	5,943.094	
Skew Angle	-	1.127	1.143	
Presence of RTLs	-	0.095	0.296	
Presence of LTLs	-	0.349	0.481	
Lighting	-	1	0	
KABCO	255	4.048	3.438	
KABC	63	1	1.666	
KAB	34	0.54	1.28	
KA	3	0.048	0.215	
SV	17	0.27	0.601	
MV	238	3.778	3.367	
Frequencies for Categorical Variables				
	Frequency Present	**Percent Present**	**Frequency Absent**	**Percent Absent**
Presence of RTLs	6	9.52	57	90.48
Presence of LTLs	22	34.92	41	65.08
Lighting	63	100	0	0.00

The major road AADTs' means and standard deviations of Florida's intersections satisfying the proposed base conditions are considerably higher than those of Ohio's major road AADTs. That is also the case with the minor road AADTs. Yet, the standard deviations of the major and minor road AADTs in both states are large. That is a factor that may inhibit transferability of Florida's SPFs of intersections conforming to proposed base conditions to Ohio. The transferability of Ohio's intersection SPFs to Florida will be deterred in a similar fashion. In addition, the skew angles' mean and standard deviation in the Florida proposed base conditions intersection data are zero while those of Ohio are considerably low. That is a factor that may facilitate SPF transferability. Yet, the proportions of turning lanes are different in each state. That may impede each state's SPFs to be transferable to the other state. The Florida and Ohio's crash rates calculated for four-leg signalized intersections that conform to the proposed base conditions are shown in Table 4.8.

Table 4.8: Florida and Ohio's Proposed Base Conditions Rural Four-Leg Signalized Intersection Crash Rates Per Hundred Million Vehicle Miles Traveled Per Year

Crash Classification	Florida	Ohio
KABCO	5.32	36.02
KABC	2.28	8.90
KAB	0.91	4.80
KA	0.15	0.42
SV	0.15	2.40
MV	5.17	33.61

As shown in Table 4.8, the crash rates of all crash classifications of Ohio are to a considerable extent larger than those of Florida. Therefore, it is expected that the SPFs for proposed base conditions of rural four-leg signalized intersections of each state are not transferable to the other state. Florida and Ohio's proposed base conditions intersection data are pooled. Their descriptive statistics are shown in Table 4.9.

Table 4.9: Florida and Ohio's Base Conditions Descriptive Statistics for Rural Four-Leg Signalized Intersections

Number of Intersections = 102	Number of Crashes	Mean	Standard Deviation	
Major Road AADT	-	9,032.907	5,713.514	
Minor Road AADT	-	3,194.676	2,820.367	
Total AADT	-	12,227.583	7,211.813	
Skew Angle	-	0.696	1.051	
Presence of RTLs	-	0.235	0.426	
Presence of LTLs	-	0.529	0.502	
Lighting	-	1	0	
KABCO	290	2.843	3.245	
KABC	78	0.765	1.394	
KAB	40	0.392	1.045	
KA	4	0.039	0.195	
SV	18	0.176	0.496	
MV	272	2.667	3.135	
Frequencies for Categorical Variables				
	Frequency Present	**Percent Present**	**Frequency Absent**	**Percent Absent**
Presence of RTLs	24	23.53	78	76.47
Presence of LTLs	54	52.94	48	47.06
Lighting	102	100.00	0	0.00

CHAPTER 5: RESEARCH METHODOLOGY

The research methodology includes developing SPFs for KABCO, KABC, KAB, KA, SV and MV crashes for each state for rural divided segments and four-leg signalized intersections. These SPFs are referred to as jurisdiction specific SPFs. Then, each state's data are pooled with those of the other states in different combinations as described in the data preparation chapter. Similarly, SPFs are developed for the same crash classifications for the pooled data. That is, two-state joint SPFs, which are models estimated from data of two states, and three-state joint SPFs, which are models estimated from data of the three states, are developed. That is performed for both average and proposed base conditions. In the following subsections, descriptions of the methods used to process the segment and intersection data are provided.

5.1 Rural Divided Multilane Highway Segments – Research Methodology

The PROC NLMIXED procedure of the statistical analysis software (SAS) version 9.4 is used for estimating models for divided segments. The KABCO, KABC, KAB, KA, SV and MV SPFs are developed for each state's average conditions segments data. The SPF takes the form of negative binomial model of the HSM for rural divided multilane segments as follows.

$$N_{SPF} = \exp[A + B\ln(AADT) + \ln(L \times y)] \tag{5.1}$$

The overdispersion parameter is defined as follows.

$$k = 1/\exp[C + D\ln(L \times y)\ ] \tag{5.2}$$

In Equations (5.1) and (5.2), L, represents the segment length and y is an offset variable which represents the number of crash years. The parameters A, B, C and D are regression coefficients. It is critical to note that the overdispersion formula in the HSM for rural divided segments' SPFs there is no coefficient multiplied by the segment length. The data are then pooled for development of joint SPFs. If two states' data are pooled the SPFs and their corresponding overdispersion parameters would take the following forms.

$$N_{SPF} = \exp[A + B\ln(AADT) + \mathrm{E}(state\ 2) + \mathrm{F}(state\ 2 \times \ln(AADT)) + \ln(L \times y)] \qquad (5.3)$$

$$k = 1/\exp[C + \mathrm{G}(state\ 2) + (D + H \times state\ 2)\ln(L \times y)\ ] \qquad (5.4)$$

The variable *state 2* is a dummy variable representing the state where crashes occurred while the other state is the reference state. In this case, four supplementary parameter coefficients *E*, *F*, *G* and *H* are estimated to capture differences in crash frequencies between states. In the case of pooling the three state's data and estimating SPFs, the joint SPF form is as follows.

$$N\_SPF = \exp[A + B\ln(AADT) + \mathrm{E}(state\ 2) + \mathrm{I}(state\ 3) + \mathrm{F}(state\ 2 \times \ln(AADT)\ ) + \mathrm{J}(state\ 3 \times \ln(AADT)\ ) + \ln(L \times y)\ ] \qquad (5.5)$$

Any state can be the reference state while the other two are identified by their dummy variables, *state 2* and *state 3*. The overdispersion parameter formula is the following.

$$k = 1/\exp[C + \mathrm{G}(state\ 2) + \mathrm{K}(state\ 3) + (D + H \times state\ 2 + M \times state\ 3)\ \ln(L \times y)\ ] \qquad (5.6)$$

The coefficients added are *I*, *J*, *K* and *M*. The analysis procedure is conducted once for average conditions and once for the proposed based conditions.

5.2 Rural Four-Leg Signalized Intersections – Research Methodology

The PROC GENMOD and NLMIXED procedures of SAS are used to develop SPFs for four-leg signalized intersections. Similar to the segments' case, each state's four-leg signalized intersection data are used for SPF development. The SPFs take the following form.

$$N_{SPF} = \exp\left[A + B\ln\left(AADT_{major}\right) + D\ln(AADT_{minor})\right] \quad (5.7)$$

The terms $AADT_{major}$ and $AADT_{minor}$ are entering AADTs from the major road and minor roads, respectively. If the transformed minor road AADT variable is statistically insignificant at the 95th percentile confidence interval, the total entering AADT, $AADT_{total}$, is used instead as shown in Equation (5.8).

$$N_{SPF} = \exp[A + J\ln(AADT_{total})] \quad (5.8)$$

The overdispersion parameter, k, is simply equal to a single value, P, to be estimated. When intersection data from two states are pooled for development of joint SPFs, the SPFs take the following form.

$$N_{SPF} = \exp\left[A + B\ln\left(AADT_{major}\right) + D\ln(AADT_{minor}) + \mathrm{E}(state\ 2) + \mathrm{F}\left(state\ 2 \times \ln\left(AADT_{major}\right)\right) + I(state\ 2 \times \ln(AADT_{minor}))\right] \quad (5.9)$$

Similarly, if the minor road AADT is insignificant at the 95th percentile confidence interval, the joint SPF takes the form as follows.

$$N_{SPF} = \exp[A + J\ln(AADT_{total}) + \mathrm{E}(state\ 2) + \mathrm{M}(state\ 2 \times \ln(AADT_{total}))] \quad (5.10)$$

Whether both the major and minor road AADTs are used or the total AADT is used, the overdispersion parameter takes the following form for the pooled SPF.

$$k = P + Q \times state\ 2 \quad (5.11)$$

The additional parameter coefficients *E*, *F*, *I*, *M* and *Q* are used to capture differences between crash frequencies in the two states of which data are used for developing the SPF. The dummy variable, *state 2*, is used to represent one state while the other state is the reference state. Both average and proposed base conditions of the four-leg signalized intersection data are processed by means of the research methodology described.

5.3 Transferability Assessment

After development of SPFs, which are based on each state's data, and joint SPFs, which are based on data pooled from multiple states, the transferability of SPFs is assessed. First, the jurisdiction specific SPFs and joint SPFs are applied to each state. Subsequently, the transferability assessment is conducted by calculating a measure called the transfer index, *TI*, which was used in similar settings by Sikder *et al*. (2014) and Hadayeghi *et al.* (2006). The *TI* value indicates the performance of the SPF being transferred to the jurisdiction of which it is being applied. The *TI* is calculated by the following formula.

$$TI = \left(LL_i(\beta_j) - LL_j(\beta_{reference\ j})\right) / \left(LL_j(\beta_j) - LL_j(\beta_{reference\ j})\right) \tag{5.12}$$

In Equation (5.12), $LL_i(\beta_j)$ is the log-likelihood of the SPF, estimated based on data, *i*, that is being implemented to predict crash frequencies in a specific jurisdiction, *j*. The terms $LL_j(\beta_j)$ and $LL_j(\beta_{reference\ j})$ are the log-likelihoods of the jurisdiction specific SPF and of the jurisdiction specific SPF with the constant only respectively. The transfer index is a measure that is used to compare the model being transferred relative to the jurisdiction specific model. High transfer index values indicate that the transferred SPF performs better than the jurisdiction specific one with the constant only. The maximum transfer index value is unity. It indicates that the transferred model is performing as well as the jurisdiction specific model. Negative transfer indices indicate that the transferred SPF underperforms the jurisdiction specific one with the constant only.

5.4 Modified Empirical Bayes Method

A modified Empirical Bayes (EB) method is proposed as a substitute to the HSM calibration factor method. That is because there is a drawback of the HSM calibration factor method. It is that every segment's predicted crash frequency is multiplied by the same calibration factor. Not all segments' or intersections' predicted crash frequencies will become nearer to the observed crash frequencies. The proposed modified EB method is used to adjust each segment or intersection's predicted crash frequency. In the modified EB method, a certain weight, w, is allocated to observed crash frequencies, N_{obs}, and predicted crash frequencies, N_{SPF}. The sum of the weighted observed and predicted crash counts is the expected crash count, N_{exp}. It is defined as shown as per the HSM.

$$N_{exp\ i} = (1 - w_i)N_{obs\ i} + w_i \times N_{SPF\ i} \tag{5.13}$$

The subscripts i are the segment or intersection number. In addition, the weight is a function of the overdispersion parameter of the SPF applied. It is defined as follows.

$$w_i = 1/(1 + k_i \times N_{SPF}) \tag{5.14}$$

Since the observed crashes for every segment or intersection for a future year is unavailable, the modified EB method is used with past observed crash data of the jurisdiction under study. That is, the term $N_{obs\ i}$ is replaced by $N_{pobs\ i}$, which represents the observed crash frequency of a past period for segment or intersection i. The basis of the modified EB method is that if the overdispersion parameter of the SPF being transferred is low, the expected crash frequency will be dependent mostly on the predicted crash frequency. On the other hand, if the overdispersion parameter is large, then the expected crash frequency will be excessively dependent on the past observed crash frequency.

5.5 Goodness of Fit Measures

The goodness of fit measures of SPFs that are used are negative twice the log-likelihood value of the SPFs, or -2*LL*, mean absolute deviation (MAD) and mean square predicted error (MSPE). The measures MAD and MSPE are defined as shown.

$$MAD = \sum_{i=1}^{N} |N_{obs\,i} - N_{SPF\,i}|/N \tag{5.15}$$

$$MSPE = \sum_{i=1}^{N} (N_{obs\,i} - N_{SPF\,i})^2/N \tag{5.16}$$

The subscripts *i*, are the segment or intersection numbers. Generally, both the MAD and MSPE are measures of the average deviation between each segment or intersection's predicted crash frequency and observed crash frequency.

CHAPTER 6: DATA ANALYSIS AND DISCUSSION FOR RURAL DIVIDED MULTILANE HIGHWAY SEGMENTS

In this chapter, results of the jurisdiction specific, two-state and three-state joint SPFs are presented and discussed. Also included are the results of the transfer indices and modified EB method. That is for both rural multilane highway segments and four-leg signalized intersections for average and proposed base conditions. In all cases, statistically insignificant variables at the 95th percentile confidence interval are removed from the SPFs and the models are re-estimated.

6.1 Rural Divided Multilane Highway Segments – Average Conditions Safety Performance Functions

The average conditions one-state and joint SPFs are developed for KABCO, KABC, KAB, KA, SV and MV crashes at segments. That is based on the 2009 through 2011 crash data of both Florida and Ohio. California's SPFs are developed using its 2009 through 2010 crash data. Hence, there are a total of 42 SPFs developed for segments conforming to the average conditions. The results of the average conditions SPFs of the segments are presented and discussed in the following subsections. They are also documented in an article accepted for presentation at the 95th annual meeting of the Transportation Research Board (Farid *et al*., 2016).

6.1.1 Jurisdiction Specific Average Conditions Segments' Safety Performance Functions

The Florida, Ohio and California jurisdiction specific SPFs are estimated successfully. Their results are shown in Tables 6.1 through 6.3.

Table 6.1: Florida's SPFs for Average Conditions Rural Divided Multilane Highway Segments

Crash Type or Severity	KABCO	KABC	KAB	KA	SV	MV
Parameters	**Parameter Estimates and P-Values**					
Constant	-8.5925 (<.0001)	-9.4251 (<.0001)	-8.4571 (<.0001)	-6.7154 (<.0001)	-3.3229 (<.0001)	-12.1531 (<.0001)
Ln(AADT)	1.0189 (<.0001)	1.0175 (<.0001)	0.8479 (<.0001)	0.5546 (<.0001)	0.3142 (<.0001)	1.3474 (<.0001)
C	0.6786 (<.0001)	0.7186 (<.0001)	0.9805 (<.0001)	0.8352 (0.0154)	0.4376 (0.0002)	0.3649 (<.0001)
D	-0.1991 (0.0003)	-0.04544 (0.5154)	-0.03383 (0.7587)	0.2077 (0.3415)	0.4729 (<.0001)	-0.2322 (<.0001)
Goodness of Fit Measures						
-2LL	7394.0	5662.7	4222.1	2326.6	3948.7	6870.6
MAD	6.263	2.678	1.355	0.581	1.122	5.766
MSPE	194.277	28.764	5.576	0.787	2.923	182.88
	- : statistically insignificant variables at alpha = 0.05 removed from the SPF					

Table 6.2: Ohio's SPFs for Average Conditions Rural Divided Multilane Highway Segments

Crash Type or Severity	KABCO	KABC	KAB	KA	SV	MV
Parameters	**Parameter Estimates and P-Values**					
Constant	-9.6493 (<.0001)	-9.5687 (<.0001)	-8.9987 (<.0001)	-8.9187 (<.0001)	-7.2307 (<.0001)	-13.934 (<.0001)
Ln(AADT)	1.0654 (<.0001)	0.9331 (<.0001)	0.8391 (<.0001)	0.6821 (<.0001)	0.7413 (<.0001)	1.4364 (<.0001)
C	0.3024 (0.0057)	0.1979 (0.4002)	0.06173 (0.8286)	-0.03457 (0.9711)	0.3091 (0.0573)	-0.3352 (0.0126)
D	0.58 (<.0001)	0.2672 (0.1559)	0.3593 (0.1257)	1.8963 (0.0591)	0.7664 (<.0001)	0.5164 (0.0003)
Goodness of Fit Measures						
-2LL	3951	2283.1	1940.2	800.4	3000.7	2691.4
MAD	1.419	0.661	0.531	0.176	0.879	0.888
MSPE	6.831	1.355	0.763	0.118	1.964	3.974
	- : statistically insignificant variables at alpha = 0.05 removed from the SPF					

Table 6.3: California's SPFs for Average Conditions Rural Divided Multilane Highway Segments

Crash Type or Severity	**KABCO**	**KABC**	**KAB**	**KA**	**SV**	**MV**
Parameters	**Parameter Estimates and P-Values**					
Constant	-9.2647 (<.0001)	-9.5299 (<.0001)	-8.4909 (<.0001)	-8.0426 (<.0001)	-8.2696 (<.0001)	-11.7955 (<.0001)
Ln(AADT)	1.0798 (<.0001)	1.0085 (<.0001)	0.8346 (<.0001)	0.6625 (<.0001)	0.8901 (<.0001)	1.2812 (<.0001)
C	0.2669 (<.0001)	0.2135 (0.0167)	0.3101 (0.0332)	0.2428 (0.5045)	0.3966 (<.0001)	-0.4304 (<.0001)
D	0.502 (<.0001)	0.7057 (<.0001)	0.9289 (<.0001)	0.4724 (0.3581)	0.426 (<.0001)	0.4834 (<.0001)
Goodness of Fit Measures						
-2LL	5721.5	3839.3	2737.2	1354.2	3889.3	4583.7
MAD	2.793	1.238	0.713	0.3	1.232	2.144
MSPE	28.477	4.869	1.244	0.253	6.549	17.169
	- : statistically insignificant variables at alpha = 0.05 removed from the SPF					

In Tables 6.1 through 6.3, the parameters representing the constants, *A*'s, are negative, indicating that for a short segment with a considerably low AADT, the sum of the constant, natural logarithm of the AADT and natural logarithm of the segment length terms will yield a value of zero or less. The exponentiation of the value results in 1 crash if the value is zero or a fraction near zero if the value is to a great extent less than zero. The coefficients for the natural logarithm of the AADT are positive and less than 1.5. For a unit increase in the natural logarithm of the AADT, the predicted crash frequency increases by the corresponding coefficient of the natural logarithm of the AADT. It should be noted that the MSPE values for KABCO and MV crashes are considerably high in Florida and California. These high values may be an indication of the presence of outliers.

6.1.2 Joint Average Conditions Segments' Safety Performance Functions

The two-state and three state SPFs for segment average conditions are also estimated successfully. In some cases, the dummy variable representing the state is statistically insignificant on the 95th percentile confidence interval while interaction terms between the dummy variables and the natural logarithm of the AADT are significant. In those cases, the interaction terms are removed instead. That is to interpret the SPFs in a more meaningful manner. The joint SPF results are shown in Tables 6.4 through 6.7.

Table 6.4: Florida and Ohio's Joint SPFs for Average Conditions Rural Divided Multilane Highway Segments

Crash Type or Severity	**KABCO**	**KABC**	**KAB**	**KA**	**SV**	**MV**
Parameters	**Parameter Estimates and P-Values**					
Constant	-8.773 (<.0001)	-9.21 (<.0001)	-8.4354 (<.0001)	-6.931 (<.0001)	-3.323 (<.0001)	-12.4247 (<.0001)
Ln(AADT)	1.0366 (<.0001)	0.9964 (<.0001)	0.8458 (<.0001)	0.5761 (<.0001)	0.3142 (<.0001)	1.3741 (<.0001)
OH	-0.6091 (<.0001)	-0.9475 (<.0001)	-0.6256 (<.0001)	-1.000 (<.0001)	-3.9077 (<.0001)	-0.9284 (<.0001)
OH× Ln(AADT)	-	-	-	-	0.4271 (<.0001)	-
C	0.6793 (<.0001)	0.7171 (<.0001)	0.9804 (<.0001)	0.8405 (0.0151)	0.4376 (0.0002)	0.3652 (<.0001)
G	-0.381 (0.0018)	-0.5039 (0.0434)	-0.917 (0.0043)	-0.9145 (0.3546)	-0.1285 (0.519)	-0.7125 (<.0001)
D	-0.1958 (0.0004)	-0.04821 (0.4897)	-0.03411 (0.7566)	0.2049 (0.3483)	0.4729 (<.0001)	-0.2278 (<.0001)
H	0.7809 (<.0001)	0.3019 (0.1324)	0.392 (0.1298)	1.6752 (0.0961)	0.2936 (0.1487)	0.7553 (<.0001)
Goodness of Fit Measures						
-2LL	11,345.0	7,946.5	6,162.3	3,127.5	6,949.4	9,562.7
MAD	3.894	1.692	0.952	0.383	1.004	3.386
MSPE	102.963	15.316	3.224	0.460	2.455	96.157
	- : statistically insignificant variables at alpha = 0.05 removed from the SPF					

As shown in Table 6.4, the dummy variable, *OH*, represents Ohio state while the reference state is Florida. The dummy variable is statistically significant and negative for all crash classifications. That indicates that if all other variable values are unchanged, Ohio experiences fewer crashes than Florida, regardless of crash classification. The interaction term between the dummy variable and natural logarithm of the AADT is significant at the 95^{th} percentile confidence interval and positive only for SV crashes. That is, the effect of the AADT on SV crashes in Ohio is more influential than in Florida. For all other crash classifications, the effects of the AADT on crash frequencies in both states are the same. The significances of the overdispersion coefficients associated with the dummy variable representing Ohio indicate that there are significant differences in crash frequencies in both states.

Table 6.5: Florida and California's Joint SPFs for Average Conditions Rural Divided Multilane Highway Segments

Crash Type or Severity	**KABCO**	**KABC**	**KAB**	**KA**	**SV**	**MV**
Parameters	**Parameter Estimates and P-Values**					
Constant	-9.0931 (<.0001)	-9.3789 (<.0001)	-8.392 (<.0001)	-7.1268 (<.0001)	-3.323 (<.0001)	-11.8394 (<.0001)
Ln(AADT)	1.0658 (<.0001)	1.013 (<.0001)	0.8415 (<.0001)	0.5955 (<.0001)	0.3142 (<.0001)	1.3166 (<.0001)
CA	-	-0.1951 (<.0001)	-0.1673 (0.0019)	-0.257 (0.0017)	-4.9467 (<.0001)	-0.3041 (<.0001)
CA×Ln(AADT)	-	-	-	-	0.5758 (<.0001)	-
C	0.6783 (<.0001)	0.7183 (<.0001)	0.9801 (<.0001)	0.8451 (0.0149)	0.4376 (0.0002)	0.3641 (<.0001)
G	-0.4231 (<.0001)	-0.5048 (<.0001)	-0.6706 (0.0012)	-0.5999 (0.2326)	-0.04098 (0.7776)	-0.7927 (<.0001)
D	-0.1727 (0.0011)	-0.04601 (0.5091)	-0.03469 (0.7526)	0.2015 (0.3564)	0.4729 (<.0001)	-0.2372 (<.0001)
H	0.6434 (<.0001)	0.751 (<.0001)	0.9613 (<.0001)	0.2812 (0.6184)	-0.04686 (0.7586)	0.7256 (<.0001)
Goodness of Fit Measures						
-2LL	13,118.0	9,502.0	6,959.3	3,681.6	7,838.0	11,455.0
MAD	4.472	1.951	1.03	0.439	1.178	3.94
MSPE	109.273	16.678	3.385	0.517	4.756	98.706
	- : statistically insignificant variables at alpha = 0.05 removed from the SPF					

For the joint Florida and California's average conditions segments' SPFs, the dummy variable representing California, *CA*, is statistically significant at the 95th percentile confidence interval and less than zero in SPFs of all crash classifications except for the KABCO SPF. The interaction term between the natural logarithm of the AADT and dummy variable is significant for the SV SPF only as is the case with the Florida and Ohio segments' joint SV SPF. The significances of the overdispersion coefficients associated with the dummy variable representing the state indicate differences in crash frequencies in Florida and California.

Table 6.6: Ohio and California's Joint SPFs for Average Conditions Rural Divided Multilane Highway Segments

Crash Type or Severity	KABCO	KABC	KAB	KA	SV	MV
Parameters	**Parameter Estimates and P-Values**					
Constant	-9.2195 (<.0001)	-9.3468 (<.0001)	-8.5024 (<.0001)	-8.091 (<.0001)	-7.8223 (<.0001)	-12.3026 (<.0001)
Ln(AADT)	1.0751 (<.0001)	0.9898 (<.0001)	0.8357 (<.0001)	0.6675 (<.0001)	0.8444 (<.0001)	1.3329 (<.0001)
OH	-0.5197 (<.0001)	-0.7498 (<.0001)	-0.465 (<.0001)	-0.6908 (<.0001)	-0.3637 (<.0001)	-0.6663 (<.0001)
OH×Ln(AADT)	-	-	-	-	-	-
C	0.2666 (<.0001)	0.2133 (0.0168)	0.31 (0.0331)	0.2422 (0.5054)	0.3913 (<.0001)	-0.4281 (<.0001)
G	0.03701 (0.7682)	-0.00157 (0.9951)	-0.2491 (0.4354)	-0.2813 (0.7821)	-0.08927 (0.6277)	0.072 (0.6279)
D	0.502 (<.0001)	0.7086 (<.0001)	0.9285 (<.0001)	0.4713 (0.3582)	0.437 (<.0001)	0.4904 (<.0001)
H	0.07615 (0.5854)	-0.4535 (0.0363)	-0.5685 (0.0513)	1.4232 (0.206)	0.3158 (0.1136)	0.04359 (0.7909)
Goodness of Fit Measures						
-2LL	9,672.5	6,122.9	4,677.3	2,154.6	6,893.2	7,276.9
MAD	2.128	0.958	0.625	0.24	1.062	1.549
MSPE	17.978	3.155	1.011	0.188	4.346	11.228
	- : statistically insignificant variables at alpha = 0.05 removed from the SPF					

For the Ohio and California average conditions segments' joint SPF, the dummy variable representing Ohio, *OH*, is statistically significant in all crash classifications and negative. Also, the interaction term between the natural logarithm of the AADT with the dummy variable is not significant in all crash types and severity levels. That is, Ohio experiences fewer crashes, regardless of classification, than California but the impact of the AADT on crashes is the same in both states' rural divided multilane highway segments. In addition, the significances of the overdispersion coefficients associated with the interaction terms capture differences in crash frequencies in Ohio and California that are not evident.

Table 6.7: Florida, Ohio and California's Joint SPFs for Average Conditions Rural Divided Multilane Highway Segments

Crash Type or Severity	KABCO	KABC	KAB	KA	SV	MV
Parameters	**Parameter Estimates and P-Values**					
Constant	-9.0923 (<.0001)	-9.2633 (<.0001)	-8.3885 (<.0001)	-7.2242 (<.0001)	-3.323 (<.0001)	-12.0694 (<.0001)
Ln(AADT)	1.0657 (<.0001)	1.0016 (<.0001)	0.8412 (<.0001)	0.6052 (<.0001)	0.3142 (<.0001)	1.3392 (<.0001)
OH	-0.5599 (<.0001)	-0.943 (<.0001)	-0.6295 (<.0001)	-0.9777 (<.0001)	-3.9077 (<.0001)	-0.9586 (<.0001)
CA	-	-0.1991 (<.0001)	-0.1674 (0.0018)	-0.2546 (0.0018)	-4.9467 (<.0001)	-0.2954 (<.0001)
OH×Ln(AADT)	-	-	-	-	0.4271 (<.0001)	-
CA×Ln(AADT)	-	-	-	-	0.5758 (<.0001)	-
C	0.6783 (<.0001)	0.7175 (<.0001)	0.9801 (<.0001)	0.8474 (0.0148)	0.4376 (0.0002)	-0.3547 (0.0072)
G	-0.3758 (0.0021)	-0.5032 (0.0438)	-0.9179 (0.0042)	-0.9091 (0.3615)	-0.1285 (0.519)	0.7194 (<.0001)
K	-0.4231 (<.0001)	-0.504 (<.0001)	-0.6706 (0.0012)	-0.6019 (0.2315)	-0.04098 (0.7776)	-0.07328 (0.6216)
D	-0.1727 (0.0011)	-0.04751 (0.4953)	-0.03473 (0.7522)	0.1996 (0.361)	0.4729 (<.0001)	0.533 (0.0002)
H	0.7527 (<.0001)	0.3001 (0.1342)	0.3936 (0.1279)	1.6858 (0.0962)	0.2936 (0.1487)	-0.7666 (<.0001)
M	0.6434 (<.0001)	0.7543 (<.0001)	0.9614 (<.0001)	0.2822 (0.6168)	-0.04686 (0.7585)	-0.04186 (0.7988)
Goodness of Fit Measures						
-2LL	17,069.0	11,786.0	8,899.4	4,482.3	10,839.0	14,148.0
MAD	3.492	1.536	0.87	0.355	1.082	2.966
MSPE	76.401	11.735	2.544	3.846	3.969	57.264
	- : statistically insignificant variables at alpha = 0.05 removed from the SPF					

In the joint three state SPFs, more variables are estimated relative to the two-state and jurisdiction specific SPFs because there are two dummy variables representing the states, interaction terms of the additional dummy variable and overdispersion coefficients for the additional dummy variable. The dummy variables are statistically significant at the 95th percentile confidence interval except for the KABCO SPF. That is, the crash frequencies in all three states are different if all variable values are the same. Yet for KABCO crashes, Florida and California's crash frequencies are not considerably different. They are also greater than the KABCO crashes in Ohio. Also, the interaction terms are statistically insignificant in all crash classifications except for SV crashes as are the cases of the two-state joint SPFs. For the three-state joint SV crashes' SPF, the significance of the interaction terms indicate that the effect of the AADT on SV crashes in California is the greatest followed by that of Ohio, which is followed by that of Florida. The significance of the overdispersion coefficients that are associated with the state dummy variables describe the differences in contributing factors in crash frequencies among all three states that are not observed.

6.1.3 General Remarks on the Joint Safety Performance Functions

The development of two and three-state joint SPFs unmistakably brings to attention the differences in evident and latent contributing factors that lead to crashes in Florida, Ohio and California. When differences in the overdispersion parameter are taken into consideration, differences in crash counts among states are captured by including dummy variables representing the states. According to the results of the analysis, it is crucial to include the state dummy parameter in the overdispersion. That is a frequently overlooked issue in SPF estimation. The coefficients belonging to the overdispersion parameter capture jurisdiction specific characteristics such as roadway geometrics, environmental factors, driver behavior factors and other factors that contribute to crashes.

6.1.4 Transferability Assessment for Segment Average Conditions

The transferability assessment is performed by applying one-state, two-state and three-state joint SPFs to a chosen jurisdiction whether Florida, Ohio or California. In addition, the transfer indices are calculated. The process is repeated for every jurisdiction and all crash classifications. Transfer indices are presented in Table 6.8.

Table 6.8: Average Conditions Rural Divided Multilane Highway Segments - Transfer Indices

Crash Type or Severity	SPF	Application Data		
		Florida	Ohio	California
KABCO	Florida	1	-0.273	0.716
	Ohio	-0.367	1	0.429
	California	0.453	0.172	1
	Florida and Ohio	1	0.999	0.72
	Florida and California	0.996	-0.033	0.998
	Ohio and California	0.454	1	1
	Florida, Ohio and California	0.995	1	0.998
KABC	Florida	1	-3.2	0.738
	Ohio	-2.718	1	-0.47
	California	0.518	-1.85	1
	Florida and Ohio	0.999	0.996	0.732
	Florida and California	1	-3.24	1
	Ohio and California	0.508	0.997	1
	Florida, Ohio and California	1	0.995	1
KAB	Florida	1	-1.573	0.728
	Ohio	-0.982	1	0.328
	California	0.453	-0.642	1
	Florida and Ohio	1	1	0.728
	Florida and California	1	-1.62	1
	Ohio and California	0.453	1	1
	Florida, Ohio and California	1	1	1
KA	Florida	1	-9.671	0.581
	Ohio	-7.464	1	-1
	California	0.352	-3.418	1
	Florida and Ohio	0.998	0.978	0.603
	Florida and California	0.994	-8.913	0.99
	Ohio and California	0.356	1	1
	Florida, Ohio and California	0.991	0.989	0.993
SV	Florida	1	0.634	-0.864
	Ohio	-6.371	1	0.512
	California	-20.822	0.187	1
	Florida and Ohio	1	1	-0.865
	Florida and California	1	0.634	1
	Ohio and California	-19.985	0.981	0.997
	Florida, Ohio and California	1	1	1
MV	Florida	1	-0.83	0.208
	Ohio	-0.818	1	0.361
	California	0.25	0.186	1
	Florida and Ohio	0.999	0.998	0.212
	Florida and California	0.999	-0.937	0.999
	Ohio and California	0.265	0.995	0.998
	Florida, Ohio and California	1	0.996	0.998

As shown in Table 6.8, the Florida SPFs being transferred to Florida state yield the maximum transfer indices of unity. This applies to all other states. Regarding the one-state SPFs, the transfer indices indicate poor transferability except for Florida and California. When two-state models are applied, the transfer indices increase considerably. For transfer indices corresponding to the two-state joint SPFs, only Ohio has negative transfer indices. Implications of the transfer index results are that Florida and Ohio regions are different. That is because each's SPFs are not transferable to the other state. Yet, when Ohio's data are combined with that of California, the SPF transferability to Florida improves. Therefore, if local data are not available, application of SPFs developed from at least two states is recommended for use. That is to prevent use of SPFs that are uncharacteristic of the actual conditions. When the three-state joint SPFs are applied, the transferability is more likely to improve because the addition of data from a state increases the sample size in the combined data. The transfer indices are near unity. It would be recommended to apply the joint three-state model to any state other than Florida, Ohio or California.

6.1.5 Modified Empirical Bayes Results for Average Conditions Segments

The modified EB method is proposed to facilitate transferability of SPFs among states. The method is applied to correct for each segment's predicted crash frequency. The proposed method is compared with the HSM calibration method. The comparison is made by applying one-state models to the other two states. Once, the models are applied to predict crash frequencies the MAD and MSPE measures are calculated for three scenarios. The first scenario is the one in which the predicted crash frequencies are kept unadjusted. In the second scenario, the predicted crash frequencies are adjusted using the HSM calibration method. The third scenario is the one in which the modified EB method is applied to correct for predicted crash frequencies. The comparisons are made for KABCO crashes and the results are shown in Table 6.9.

Table 6.9: Comparison of Predicted, Calibrated and Expected KABCO Crash Frequencies – Average Conditions Rural Divided Multilane Highway Segments

Goodness of Fit Measures	Based on Predicted Crash Frequency	Based on Predicted Crash Frequency Multiplied by the HSM's Calibration Factor	Based on Predicted Crash Frequency with EB Correction
	Application of Florida's KABCO SPF to Ohio		
MAD	2.329	1.423	0.676
MSPE	15.194	6.883	0.663
	Application of Florida's KABCO SPF to California		
MAD	2.906	2.679	0.890
MSPE	29.550	26.949	2.187
	Application of Ohio's KABCO SPF to Florida		
MAD	4.609	5.149	1.075
MSPE	161.521	163.573	4.184
	Application of Ohio's KABCO SPF to California		
MAD	2.471	2.676	0.848
MSPE	28.418	27.076	2.233
	Application of California's KABCO SPF to Florida		
MAD	5.977	5.142	0.820
MSPE	186.324	163.869	1.669
	Application of California's KABCO SPF to Ohio		
MAD	2.088	1.423	0.629
MSPE	12.386	6.816	0.649

As shown in Table 6.9, the HSM calibration method does not necessarily improve the accuracy of the prediction of the SPFs. For instance, when Ohio's KABCO SPF is applied to Florida and the HSM calibration factor is multiplied by the predicted crash frequencies, the resulting prediction becomes less accurate. On the other hand, the modified EB method markedly reduces the MADs and MSPEs relative to the HSM calibration method. Therefore, the modified EB method is convenient for adjusting crash counts predicted by the transferred SPFs.

6.2 Rural Divided Multilane Highway Segments – Modified Base Conditions' Safety Performance Functions

The rural divided highway segments average conditions data are subset to include only segments that satisfy the modified base conditions. The Florida and Ohio one-state SPFs are developed using each state's crash data from the years 2009 to 2011. California's SPFs are developed from the 2009 and 2011 crash data. Furthermore, only the Ohio and California data are pooled for development of joint SPFs. That is because of the lack of segments satisfying the modified base conditions in Florida relative to the other states. The analysis results for the segments' modified base conditions are presented and discussed in the following subsections.

6.2.1 Jurisdiction Specific Modified Base Conditions Segments' Safety Performance Functions

The jurisdiction specific SPFs for divided segments that conform to the modified base conditions are developed successfully. Their results are shown in Tables 6.10 through 6.12.

Table 6.10: Florida's SPFs for Modified Base Conditions Rural Divided Multilane Highway Segments

Crash Type or Severity	KABCO	KABC	KAB	KA	SV	MV
Parameters	**Parameter Estimates and P-Values**					
Constant	-4.5044 (0.0207)	-6.4534 (0.0335)	-6.2837 (0.0651)	-4.5994 (0.3393)	-2.1744 (0.3665)	-7.1909 (0.012)
Ln(AADT)	0.5518 (0.0038)	0.6776 (0.0221)	0.6134 (0.0622)	0.3654 (0.4315)	0.2195 (0.3461)	0.7732 (0.0058)
C	1.4488 (0.0016)	0.9281 (0.0589)	0.608 (0.3002)	-0.2127 (0.7795)	12.5049 (0.9061)	0.7596 (0.0469)
D	1.0254 (0.0778)	0.9925 (0.1041)	1.1196 (0.0779)	1.2081 (0.0618)	0.6444 (0.9839)	0.9318 (0.0778)
Goodness of Fit Measures						
-2LL	274.0	218.6	174.4	106.4	158.6	248.5
MAD	2.083	1.603	1.137	0.607	0.885	1.917
MSPE	8.927	5.073	2.630	0.801	1.352	7.957
	- : statistically insignificant variables at alpha = 0.05 removed from the SPF					

Table 6.11: Ohio's SPFs for Modified Base Conditions Rural Divided Multilane Highway Segments

Crash Type or Severity	KABCO	KABC	KAB	KA	SV	MV
Parameters	**Parameter Estimates and P-Values**					
Constant	-11.6144 (<.0001)	-13.2993 (<.0001)	-14.1526 (<.0001)	-12.9638 (0.0006)	-9.7614 (<.0001)	-16.0015 (<.0001)
Ln(AADT)	1.2629 (<.0001)	1.3075 (<.0001)	1.3666 (<.0001)	1.092 (0.0063)	1.0119 (<.0001)	1.6287 (<.0001)
C	0.5562 (0.0066)	0.1494 (0.7333)	0.08201 (0.8776)	11.1799 (0.962)	0.2857 (0.2632)	-0.1121 (0.6614)
D	0.6307 (0.0037)	0.5468 (0.1253)	0.3836 (0.3446)	0.7514 (0.991)	0.8936 (0.0009)	0.7035 (0.0125)
Goodness of Fit Measures						
-2LL	1413	751.8	641.7	263.2	1124.8	929.4
MAD	1.366	0.61	0.507	0.166	0.986	0.756
MSPE	4.194	0.929	0.686	0.091	1.963	1.771
	- : statistically insignificant variables at alpha = 0.05 removed from the SPF					

Table 6.12: California's SPFs for Modified Base Conditions Rural Divided Multilane Highway Segments

Crash Type or Severity	KABCO	KABC	KAB	KA	SV	MV
Parameters	**Parameter Estimates and P-Values**					
Constant	-9.8162 (<.0001)	-10.0834 (<.0001)	-9.7057 (<.0001)	-8.6512 (<.0001)	-7.8579 (<.0001)	-13.6728 (<.0001)
Ln(AADT)	1.1055 (<.0001)	1.0357 (<.0001)	0.9328 (<.0001)	0.704 (<.0001)	0.8308 (<.0001)	1.4265 (<.0001)
C	0.4915 (<.0001)	0.4047 (0.0142)	0.5033 (0.0836)	-0.00616 (0.9925)	1.0947 (<.0001)	-0.3918 (0.0005)
D	0.5836 (<.0001)	0.663 (0.0007)	1.0829 (0.0009)	1.6728 (0.0622)	0.4383 (0.0597)	0.6527 (<.0001)
Goodness of Fit Measures						
-2LL	2258.6	1495.5	1087.9	551.7	1617.0	1689.6
MAD	2.285	1.058	0.648	0.287	1.064	1.776
MSPE	19.850	3.070	0.915	0.242	3.074	15.608
	- : statistically insignificant variables at alpha = 0.05 removed from the SPF					

As shown in Tables 6.10 through 6.12, the natural logarithm of the AADT variable is statistically insignificant at the 95^{th} percentile confidence interval for Florida's KAB, KA and SV crashes. That is because of the low sample size of 57 segments. There is a lack of data to support the assertion that the AADT influences non-incapacitating, incapacitating and single-vehicle crashes. That is unlike the cases of Ohio and California.

6.2.2 Ohio and California Joint Modified Base Conditions Divided Segments' Safety Performance Functions

The Ohio and California data of the divided segments conforming to the base conditions are combined. The developed joint SPFs of the pooled data are shown in Table 6.13.

Table 6.13: Ohio and California's Joint SPFs for Modified Base Conditions Rural Divided Multilane Highway Segments

Crash Type or Severity	**KABCO**	**KABC**	**KAB**	**KA**	**SV**	**MV**
Parameters	**Parameter Estimates and P-Values**					
Constant	-10.4375 (<.0001)	-11.0919 (<.0001)	-10.6220 (<.0001)	-9.8039 (<.0001)	-8.3530 (<.0001)	-14.5149 (<.0001)
Ln(AADT)	1.1371 (<.0001)	1.0723 (<.0001)	0.9904 (<.0001)	0.7557 (<.0001)	0.8613 (<.0001)	1.4704 (<.0001)
CA	0.3096 (<.0001)	0.6451 (<.0001)	0.3416 (0.0046)	0.6377 (0.0011)	0.1922 (0.0151)	0.4072 (0.0002)
CA×Ln(AADT)	-	-	-	-	-	-
C	0.5442 (0.0073)	0.1215 (0.7788)	0.05130 (0.9225)	9.7495 (0.9715)	0.2823 (0.2693)	-0.1336 (0.5964)
G	-0.05102 (0.8236)	0.2822 (0.5422)	0.4394 (0.4653)	-9.7658 (0.9715)	0.8175 (0.0135)	-0.2578 (0.3501)
D	0.6473 (0.0027)	0.5680 (0.1079)	0.4024 (0.3237)	1.0930 (0.9937)	0.8964 (0.0008)	0.7255 (0.0094)
H	-0.06195 (0.8018)	0.09297 (0.8177)	0.6697 (0.1977)	0.5744 (0.9967)	-0.4491 (0.2051)	-0.07034 (0.8177)
Goodness of Fit Measures						
-2LL	3672.7	2248.8	1732.3	815.7	2743.1	2619.9
MAD	1.901	0.868	0.588	0.235	1.031	1.347
MSPE	13.149	2.161	0.816	0.177	2.581	9.794
	- : statistically insignificant variables at alpha = 0.05 removed from the SPF					

As shown in Table 6.13, the dummy variables representing California, *CA*, are statistically significant at the 95th percentile confidence interval in the SPFs of all crash classifications. The dummy variables are also positive indicating that with the AADT being unchanged, California's segments conforming to the modified base conditions experience more KABCO, KABC, KAB, KA, SV and MV crashes than Ohio's. In the case of the Ohio and California's average conditions divided segments' joint SPFs, the dummy variable representing California is also significant in all crash classifications but negative. That is, either the lane width, shoulder width, median width, lighting conditions or automated speed enforcement contributes significantly to crashes. It should be noted that the KA crashes' SPF is run with the interaction term between the dummy variable representing the state with the natural logarithm of the AADT and did not converge. Removal of the interaction term caused the model to converge. In addition, the statistical insignificance of the interaction terms indicate that the effect of the AADT on crashes, regardless of type or severity level, in California is not different from that in Ohio. That is similar to the case of Ohio and California's joint SPFs for divided segments conforming to average conditions.

6.2.3 Transferability Assessment for Divided Segments Modified Base Conditions

The transferability assessment is conducted by applying the modified base conditions jurisdiction specific SPFs to each state and calculating the transfer indices. Also, the Ohio and California's modified base conditions segments' joint SPFs are applied to each state and the transfer indices are calculated. Transfer index results are shown in Table 6.14. As previously stated in the data preparation chapter, the Florida data are not pooled with those of any other state because the number of segments, conforming to the modified base conditions, in Florida is low relative to those of the other two states. That may bias the results.

Table 6.14: Modified Base Conditions Rural Divided Multilane Highway Segments - Transfer Indices

Crash Type or Severity	SPF	Application Data		
		Florida	Ohio	California
KABCO	Florida	1	-1.478	-0.076
	Ohio	-0.948	1	0.838
	California	-1.229	0.505	1
	Ohio and California	-0.485	0.990	0.999
KABC	Florida	1	-5.705	0.726
	Ohio	-1.009	1	0.339
	California	0.649	-1.418	1
	Ohio and California	-2.405	0.968	0.999
KAB	Florida	1	-3.803	0.654
	Ohio	-0.359	1	0.683
	California	0.587	-0.112	1
	Ohio and California	-1.825	0.928	0.996
KA	Florida	1	-20.978	-1.803
	Ohio	-38.673	1	-0.676
	California	-11.185	-2.426	1
	Ohio and California	-59.267	0.914	0.994
SV	Florida	1	-1.290	-1.544
	Ohio	NA	1	0.710
	California	NA	0.451	1
	Ohio and California	NA	0.978	0.998
MV	Florida	1	-2.481	-0.094
	Ohio	-0.441	1	0.779
	California	-0.843	0.508	1
	Ohio and California	-0.436	0.990	0.999
	NA : Constants only model of jurisdiction specific data failed to converge			

The Florida KABC and KAB SPFs are transferable to California and vice versa as indicated by the high transfer indices. They are even higher than their counterparts in the divided segments' average conditions. That is because both states' divided segments' characteristics conform to the modified base conditions. Yet, the transfer indices of Florida to California and vice versa for

KABCO crashes for modified base conditions are lower than those of the average conditions. That is most likely because of the low sample size of segments conforming to the modified base conditions in Florida. On the other hand, Florida's SPFs, regardless of crash classification, are not transferable to Ohio and vice versa as indicated by the negative transfer indices. Furthermore, it crucial to note that for SV crashes, the constants only SPF of Florida failed to converge inhibiting calculation of transfer indices of the other two states' SV SPFs and the joint SV SPF to Florida state. Finally, the Ohio and California joint SPFs are not transferable to Florida as indicated by the negative transfer indices as opposed to the case of the average conditions. Similarly, that is possibly because there are only 57 divided segments satisfying the modified base conditions in Florida.

CHAPTER 7: DATA ANALYSIS AND DISCUSSION FOR RURAL FOUR-LEG SIGNALIZED INTERSECTIONS

7.1 Rural Four-Leg Signalized Intersections – Average Conditions Safety Performance Functions

The analysis procedure is different for four-leg signalized intersections. Florida and Ohio's average conditions four-leg signalized intersection data from 2009 to 2010 are used for development of one-state and joint SPFs. California's 2009 data are also used for estimating jurisdiction specific SPFs. Florida's 2011 data, Ohio's 2011 data and California's 2010 data are used for application of the modified EB method. The results of the average conditions SPFs for four-leg signalized intersections are presented and discussed in the following subsections. Since the number of intersections sampled in California is limited, only the Florida and Ohio data are used for estimation of joint SPFs.

7.1.1 Jurisdiction Specific Average Conditions Rural Four-Leg Signalized Intersections' Safety Performance Functions

The jurisdiction specific SPFs for average conditions are developed successfully except for SV crashes in Florida and KA crashes in California. That is mainly because there are only 7 SV crashes in Florida and 4 KA crashes in California. The jurisdiction specific SPF results are shown in Tables 7.1 through 7.3.

Table 7.1: Florida's SPFs for Average Conditions Rural Four-Leg Signalized Intersections

Crash Type or Severity	KABCO	KABC	KAB	KA	SV	MV
Parameters	**Parameter Estimates and P-Values**					
Constant	-10.4409 (0.0003)	-9.5045 (0.0047)	-9.0437 (0.0228)	-16.2787 (0.0274)	Fail to converge	-11.0295 (0.0002)
Ln(AADT major)	NA	NA	NA	NA		NA
Ln(AADT minor)	NA	NA	NA	NA		NA
Ln(AADT total)	1.0049 (0.0006)	0.8449 (0.0145)	0.7227 (0.0758)	1.3346 (0.0737)		1.0596 (0.0006)
k	1.349	1.3997	0.5891	0		1.4469
Goodness of Fit Measures						
-2LL	351.8	254.2	164.6	68.0	Fail to converge	340.9
MAD	1.07	0.68	0.395	0.136		1.03
MSPE	2.238	0.912	0.291	0.068		2.129
	NA : not applicable variables					

Table 7.2: Ohio's SPFs for Average Conditions Rural Four-Leg Signalized Intersections

Crash Type or Severity	KABCO	KABC	KAB	KA	SV	MV
Parameters	**Parameter Estimates and P-Values**					
Constant	-5.1407 (<.0001)	-9.5198 (<.0001)	-9.5693 (<.0001)	-5.9594 (0.1902)	0.6957 (0.8308)	-5.7121 (<.0001)
Ln(AADT major)	0.5131 (<.0001)	0.8321 (<.0001)	0.7486 (0.0005)	NA	NA	0.5736 (<.0001)
Ln(AADT minor)	0.1728 (0.0006)	0.2082 (0.032)	0.2474 (0.0396)	NA	NA	0.1675 (0.0011)
Ln(AADT total)	NA	NA	NA	0.3592 (0.4594)	-0.3108 (0.3801)	NA
k	0.2293	0.8209	1.128	4.3726	1.2686	0.239
Goodness of Fit Measures						
-2LL	566.6	371.5	288.5	103.4	143.5	556.1
MAD	2.48	1.354	0.901	0.262	0.376	2.407
MSPE	13.378	5.041	2.28	0.239	0.29	12.788
	NA : not applicable variables					

Table 7.3: California's SPFs for Average Conditions Rural Four-Leg Signalized Intersections

Crash Type or Severity	KABCO	KABC	KAB	KA	SV	MV
Parameters	**Parameter Estimates and P-Values**					
Constant	-7.6211 (0.0052)	-5.0492 (0.1270)	-5.0546 (0.5639)	Fail to converge	5.2161 (0.4895)	-9.1584 (0.0005)
Ln(AADT major)	NA	NA	NA		NA	0.8594 (0.0006)
Ln(AADT minor)	NA	NA	NA		NA	0.1992 (0.0344)
Ln(AADT total)	0.8605 (0.0012)	0.5308 (0.1024)	0.4415 (0.6075)		-0.6166 (0.4136)	NA
k	0.1566	4.04×10^{-6}	1.251×10^{-6}		1.3221	0.0784
Goodness of Fit Measures						
-2LL	136.7	94.8	65.8	Fail to converge	54.6	124.7
MAD	1.756	0.804	0.613		0.554	1.451
MSPE	5.502	1.109	0.531		0.58	3.791
	NA : not applicable variables					

As shown in Tables 7.1 through 7.3, the natural logarithm of the total entering AADT is used instead of those of the major and minor road AADTs in several SPFs. That is because the transformed minor road AADT variables are statistically insignificant at the 95th percentile confidence interval. It is critical to note that while Florida's SV crashes' SPF fails to converge, in those of Ohio and California, the constant terms are positive implying abnormally high SV crash frequencies even for low AADTs. Also, the natural logarithm of the total AADTs are insignificant. That is an indication that the total entering AADT, an exposure measure, does not influence SV crashes. The positive constant and insignificant transformed AADT variable are results that are inconsistent with those of the segments. It is most likely because of the limited number of SV crashes in Ohio and California. The observed SV crash frequencies are 28 and 13 in Ohio and California, respectively.

7.1.2 Florida and Ohio's Joint Average Conditions Rural Four-Leg Signalized Intersections' Safety Performance Functions

The Florida and Ohio joint average conditions four-leg signalized intersections' SPFs are run successfully for all crash classifications except for KA crashes. The KA model failed to converge. The results of the joint SPFs are shown in Table 7.4.

Table 7.4: Florida and Ohio's Joint SPFs for Average Conditions Rural Four-Leg Signalized Intersections

Crash Type or Severity	KABCO	KABC	KAB	KA	SV	MV
Parameters	**Parameter Estimates and P-Values**					
Constant	-7.1980 (<.0001)	-8.5937 (<.0001)	-8.3259 (0.0004)	Failed to converge	-1.4718 (0.6333)	-7.7673 (<.0001)
Ln(AADT major)	0.5364 (<.0001)	NA	NA		NA	0.5923 (<.0001)
Ln(AADT minor)	0.1799 (0.0003)	NA	NA		NA	0.1785 (0.0004)
Ln(AADT total)	NA	0.8437 (<.0001)	0.7530 (0.0023)		-0.2237 (0.4853)	NA
OH	1.7926 (<.0001)	-	-		1.3672 (0.0025)	1.8035 (<.0001)
OH ×Ln (AADT major)	-	NA	NA		NA	-
OH ×Ln (AADT minor)	-	NA	NA		NA	-
OH ×Ln (AADT total)	NA	-	-		-	NA
P	1.4092 (0.0005)	3.1814 (0.0015)	3.8424 (0.0366)		2.596×10^{-6} (-)	1.5040 (0.0007)
Q	-1.1796 (0.0039)	-2.2110 (0.0391)	-2.4381 (0.2242)		1.1819 (0.2292)	-1.2646 (0.0045)
Goodness of Fit Measures						
-2LL	920.2	658.2	484.9	Failed to converge	199.0	898.6
MAD	1.770	1.158	0.759		0.234	1.715
MSPE	7.668	3.349	1.430		0.166	7.328
	NA : not applicable variables - : statistically insignificant variables at alpha = 0.05 removed from the SPF (-) : p-value cannot be computed due to low overdispersion parameter					

As shown in Table 7.4, the interaction terms between the AADT, whether from the major road, minor road or both, and the dummy variable representing the state are statistically insignificant at the 95th percentile confidence interval. That is for all crash types and severity levels. Therefore, the effect of the AADTs on crashes in Ohio is not different than that in Florida. For KABC and KAB crashes, the dummy variable representing Ohio, *OH*, is statistically insignificant indicating that for the same total AADTs, KABC and KAB crash frequencies in Florida and Ohio are the same. That is unlike the case of average conditions of divided segments where Florida and Ohio's crash patterns are significantly different for all crash classifications given the same AADT.

7.1.3 Transferability Assessment for Four-Leg Signalized Intersections' Average Conditions

The transfer indices are calculated by applying the one-state and joint SPFs to each jurisdiction. The model log-likelihoods are compared as well by calculating the transfer indices. Transfer index results are shown in Table 7.5.

Table 7.5: Average Conditions Rural Four-Leg Signalized Intersections - Transfer Indices

Crash Type or Severity	SPF	Application Data		
		Florida	Ohio	California
KABCO	Florida	1.000	-9.723	-5.963
	Ohio	-39.465	1.000	0.431
	California	-27.737	-0.096	1.000
	Florida and Ohio	-1.118	0.998	-12.637
KABC	Florida	1.000	-2.340	-4.787
	Ohio	-14.305	1.000	-0.414
	California	-40.740	-2.475	1.000
	Florida and Ohio	-5.585	0.658	-0.899
KAB	Florida	1.000	-5.401	-23.393
	Ohio	-22.081	1.000	-9.226
	California	-25.931	-2.012	1.000
	Florida and Ohio	-8.068	0.502	-6.016
KA	Florida	1.000	NA	NA
	Ohio	-1.998	1.000	NA
	California	-6.587	-24.423	NA
	Florida and Ohio	NA	NA	NA
SV	Florida	NA	NA	NA
	Ohio	NA	1.000	-21.431
	California	NA	-98.999	1.000
	Florida and Ohio	NA	0.910	-63.004
MV	Florida	1.000	-8.497	-3.094
	Ohio	-37.312	1.000	0.385
	California	-25.912	-2.183	1.000
	Florida and Ohio	-1.056	0.998	-6.055
	NA : Transfer index not calculated because of corresponding SPF failure to converge			

The KA models' failure to converge deters calculation of transfer indices as shown in Table 7.5. Generally, the transfer indices indicate that neither jurisdiction specific SPFs are transferable to any other jurisdiction. Also, the joint Florida and Ohio SPFs are not transferable to any state. This is unlike the average conditions segments case where Florida's SPFs are transferable to California and vice versa except for single vehicle crashes' SPFs.

7.1.4 Application of the Modified Empirical Bayes Method to Average Conditions Rural Four-Leg Signalized Intersections

As previously stated, the 2011 data of Florida, 2011 data of Ohio and 2010 data of California are kept for application of the modified EB method. These data are termed validation datasets. First, the developed SPFs are applied to the validation data to obtain the predicted crash frequencies, denoted by N_{SPF}. Each SPF is applied to all jurisdictions for all crash classifications. The next step is the calculation of the expected crash frequency, N_{exp}, per intersection which is a function of not only the predicted crash frequency but also of the past observed one. Since the SPFs are used to predict crash frequencies per year, the past observed crash frequencies are averaged over the years for every intersection to yield the past observed crash frequency per year, N_{pobs}, to be applied in the EB equation. For instance, for an intersection i in Florida, the past observed crash frequency is the average of the observed crash frequencies in 2009 and 2010. That is the case if the transferred SPFs were applied to the Florida jurisdiction. The same applies for Ohio. For California, the past observed crash frequencies are simply those of 2009. Once the average observed crash frequencies are obtained, the expected crash frequencies are calculated. Then, the MAD and MSPE measures are calculated for three settings. In the first setting, the unadjusted predicted crash frequencies, N_{SPF}, are compared with the observed ones of the validation data. The second setting is the one where the predicted crash frequencies are calibrated by means of the HSM method using the average of the past observed crash frequencies. The calibrated predictions are compared with the observed crash frequencies of the validation data. Finally, in the third setting, the expected crash frequencies are compared with the observed crash counts of the validation data. That is to compare the performance of the modified EB method with that of the HSM calibration method. Results of the MAD and MSPE are presented in Tables 7.6 through 7.8.

Table 7.6: Comparison of Predicted, Calibrated and Expected Crash Frequencies in Florida – Average Conditions Rural Four-Leg Signalized Intersections

Crash Type or Severity	SPF	Application Data: Florida 2011					
		MAD	MSPE	MAD Calculated Using HSM Calibration Factor	MSPE Calculated Using HSM Calibration Factor	MAD Calculated Using Modified EB	MSPE Calculated Using Modified EB
KABCO	Florida	2.997	20.346	2.995	20.325	2.975	19.526
	Ohio	2.307	10.669	3.013	20.587	2.396	12.882
	California	2.306	12.346	3.003	20.446	2.428	13.722
	Florida and Ohio	3.019	20.644	3.011	20.561	2.986	19.679
KABC	Florida	1.707	6.856	1.707	6.853	1.693	6.717
	Ohio	1.43	4.567	1.699	6.813	1.554	5.441
	California	1.476	4.734	1.721	6.928	1.476	4.734
	Florida and Ohio	1.555	5.534	1.707	6.853	1.645	6.212
KAB	Florida	0.918	1.93	0.918	1.93	0.916	1.926
	Ohio	0.823	1.262	0.912	1.917	0.867	1.5
	California	0.887	1.478	0.924	1.943	0.887	1.478
	Florida and Ohio	0.879	1.56	0.917	1.929	0.892	1.76
KA	Florida	0.32	0.402	0.32	0.402	0.32	0.402
	Ohio	0.344	0.381	0.323	0.404	0.343	0.394
	California	NA	NA	NA	NA	NA	NA
	Florida and Ohio	NA	NA	NA	NA	NA	NA
SV	Florida	NA	NA	NA	NA	NA	NA
	Ohio	0.24	0.162	0.187	0.176	0.233	0.163
	California	0.54	0.325	0.187	0.176	0.386	0.188
	Florida and Ohio	0.187	0.175	0.187	0.175	0.187	0.175
MV	Florida	2.861	18.971	2.86	18.96	2.844	18.235
	Ohio	2.268	10.175	2.88	19.2	2.309	12.192
	California	2.266	12.264	2.862	18.978	2.322	12.883
	Florida and Ohio	2.885	19.254	2.878	19.176	2.856	18.378
	NA : Measure not calculated because of corresponding SPF failure to converge						

Table 7.7: Comparison of Predicted, Calibrated and Expected Crash Frequencies in Ohio – Average Conditions Rural Four-Leg Signalized Intersections

Crash Type or Severity	SPF	Application Data: Ohio 2011					
		MAD	MSPE	MAD Calculated Using HSM Calibration Factor	MSPE Calculated Using HSM Calibration Factor	MAD Calculated Using Modified EB	MSPE Calculated Using Modified EB
KABCO	Florida	2.343	9.837	1.329	3.857	1.627	5.146
	Ohio	1.267	3.656	1.268	3.575	1.099	2.714
	California	1.366	4.617	1.294	3.674	1.208	3.494
	Florida and Ohio	1.267	3.642	1.265	3.58	1.102	2.702
KABC	Florida	1.001	2.339	0.913	1.543	0.892	1.831
	Ohio	0.9	1.551	0.898	1.531	0.868	1.454
	California	0.942	1.55	0.934	1.605	0.942	1.55
	Florida and Ohio	0.938	1.783	0.913	1.543	0.863	1.515
KAB	Florida	0.696	1.409	0.732	1.046	0.682	1.308
	Ohio	0.726	1.051	0.729	1.04	0.653	0.874
	California	0.744	1.122	0.755	1.076	0.744	1.122
	Florida and Ohio	0.713	1.174	0.731	1.044	0.65	0.949
KA	Florida	0.156	0.127	0.193	0.12	0.156	0.127
	Ohio	0.19	0.119	0.195	0.118	0.185	0.117
	California	NA	NA	NA	NA	NA	NA
	Florida and Ohio	NA	NA	NA	NA	NA	NA
SV	Florida	NA	NA	NA	NA	NA	NA
	Ohio	0.227	0.17	0.249	0.168	0.228	0.17
	California	0.585	0.406	0.247	0.168	0.427	0.249
	Florida and Ohio	0.228	0.17	0.25	0.168	0.23	0.17
MV	Florida	2.234	9.049	1.289	3.53	1.55	4.614
	Ohio	1.249	3.341	1.251	3.285	1.064	2.421
	California	1.495	4.901	1.292	3.691	1.389	4.154
	Florida and Ohio	1.249	3.332	1.25	3.292	1.066	2.41
	NA : Measure not calculated because of corresponding SPF failure to converge						

Table 7.8: Comparison of Predicted, Calibrated and Expected Crash Frequencies in California – Average Conditions Rural Four-Leg Signalized Intersections

Crash Type or Severity	SPF	Application Data: California 2010					
		MAD	MSPE	MAD Calculated Using HSM Calibration Factor	MSPE Calculated Using HSM Calibration Factor	MAD Calculated Using Modified EB	MSPE Calculated Using Modified EB
KABCO	Florida	2.193	9.468	1.958	5.965	2.118	8.377
	Ohio	2.148	5.956	1.784	4.922	1.932	6.652
	California	1.973	5.788	1.88	5.54	1.933	6.618
	Florida and Ohio	2.383	10.201	1.795	4.974	2.059	7.893
KABC	Florida	1.22	2.989	1.052	2.137	1.094	2.594
	Ohio	1.157	2.384	1.085	2.238	1.214	2.317
	California	1.031	1.961	1.007	1.962	1.031	1.961
	Florida and Ohio	0.985	2.213	1.051	2.137	1.256	2.654
KAB	Florida	0.56	0.734	0.696	0.68	0.595	0.729
	Ohio	0.82	0.939	0.713	0.723	0.816	0.865
	California	0.676	0.645	0.691	0.655	0.676	0.645
	Florida and Ohio	0.668	0.661	0.697	0.684	0.812	0.991
KA	Florida	0.168	0.101	0.224	0.102	0.168	0.101
	Ohio	0.19	0.1	0.23	0.101	0.214	0.124
	California	NA	NA	NA	NA	NA	NA
	Florida and Ohio	NA	NA	NA	NA	NA	NA
SV	Florida	NA	NA	NA	NA	NA	NA
	Ohio	0.223	0.214	0.337	0.221	0.238	0.221
	California	0.458	0.291	0.343	0.232	0.423	0.287
	Florida and Ohio	0.173	0.225	0.335	0.219	0.173	0.225
MV	Florida	2.059	8.508	1.856	5.676	1.985	7.642
	Ohio	2.134	6.075	1.669	4.615	1.896	6.251
	California	1.832	5.469	1.794	5.362	1.813	5.812
	Florida and Ohio	2.241	9.107	1.68	4.666	1.94	7.239
	NA : Measure not calculated because of corresponding SPF failure to converge						

The MADs and MSPEs are not calculated for California's SPF for KA crashes, the joint SPF for KA crashes and Florida's SPF for SV crashes since the corresponding models failed to converge. When the SPFs are applied to Florida's validation data, the results indicate that the modified EB method outperforms the HSM calibration method except for SV and KA crashes. That is possibly due to lack of SV and KA crashes in the data based on which the corresponding SPFs are developed. When the SPFs are applied to the Ohio validation data, the MADs and MSPEs obtained from the modified EB method are lower relative to those obtained from the HSM calibration method. Exceptions are Florida's KABCO crashes' SPF, California's KABC crashes' SPF, California's SV crashes' SPF, Florida's MV crashes' SPF and California's MV crashes' SPF. When the California data are used for validation, the HSM calibration method outperforms the modified EB method. It is crucial to note that the EB method does not necessarily outperform the HSM calibration method for every SPF especially for intersections. That is because the EB weight is a function of the overdispersion parameter which is the same for every intersection. In the case of the divided segments, the overdispersion parameter varies from segment to segment. Also, it is not recommended to apply either the HSM method or the modified EB method to adjust for crash predictions that are in the distant future.

7.2 Rural Four-Leg Signalized Intersections – Proposed Base Conditions' Safety Performance Functions

The data of the four-leg signalized intersections conforming to the base conditions are processed the same way as those of the average conditions. As previously stated, the proposed base conditions are: street lighting presence, skew angles within the range of 0^{o} to 5^{o} and no automated speed enforcement. The condition of no turning lanes is not proposed since there is an inadequate number of intersections in all three states without turning lanes. The 2009 and 2010 modified base conditions four-leg signalized intersection data of both Florida and Ohio are analyzed. The 2011 data of both states are kept for validation. However California's four-leg intersection datasets are not processed since no data are available regarding the skew angles. Therefore, SPFs are developed for four-leg signalized intersections conforming to the proposed base conditions from the Florida data, Ohio data and the pooled data of both states. The SPF results are shown and discussed in the following subsections.

7.2.1 Jurisdiction Specific Proposed Base Conditions Rural Four-Leg Signalized Intersections' Safety Performance Functions

The jurisdiction specific SPFs are estimated for Florida and Ohio. The results are presented in Tables 7.9 through 7.10.

Table 7.9: Florida's SPFs for Proposed Base Conditions Rural Four-Leg Signalized Intersections

Crash Type or Severity	**KABCO**	**KABC**	**KAB**	**KA**	**SV**	**MV**
Parameters	**Parameter Estimates and P-Values**					
Constant	-15.1707 (0.0088)	Fail to converge	Fail to converge	Fail to converge	23.7093 (0.359)	-17.3419 (0.0032)
Ln(AADT major)	NA				NA	NA
Ln(AADT minor)	NA				NA	NA
Ln(AADT total)	1.4801 (0.0132)				-3.0647 (0.2917)	1.6977 (0.005)
k	1.8485				0	1.5029
Goodness of Fit Measures						
-2LL	91.6	Fail to converge	Fail to converge	Fail to converge	7.7	88.9
MAD	0.907				0.048	0.851
MSPE	1.853				0.025	1.69
	NA : not applicable variables					

Table 7.10: Ohio's SPFs for Proposed Base Conditions Rural Four-Leg Signalized Intersections

Crash Type or Severity	**KABCO**	**KABC**	**KAB**	**KA**	**SV**	**MV**
Parameters	**Parameter Estimates and P-Values**					
Constant	-4.9223 (0.0002)	-11.1361 (0.0002)	-15.7254 (0.0001)	Fail to converge	2.2938 (0.5956)	-5.9217 (<.0001)
Ln(AADT major)	0.4941 (0.0005)	NA	NA		NA	0.6059 (<.0001)
Ln(AADT minor)	0.171 (0.0107)	NA	NA		NA	0.1619 (0.0179)
Ln(AADT total)	NA	1.1285 (0.0004)	1.5458 (0.0004)		-0.4776 (0.3213)	NA
k	0.2067	0.6414	0.8953		1.1426	0.2139
Goodness of Fit Measures						
-2LL	282.1	161.9	110.7	Fail to converge	81.7	274.2
MAD	2.271	0.946	0.623		0.428	2.175
MSPE	9.467	2.301	1.321		0.351	9.077
	NA : not applicable variables					

The KABC, KAB and KA crashes' SPFs of Florida failed to converge mainly because of the low crash frequencies. For SV crashes, the SPF converges but the constant is large while the natural logarithm of the total AADT is statistically insignificant on the 95th percentile confidence interval. Similarly that is due to the limited number of SV crashes. There are records of 15 KABC crashes, 6 KAB crashes, 1 KA crash and 1 SV crash in Florida's four-leg signalized intersections satisfying the proposed base conditions. For Ohio's case, the KA crashes' SPF failed to converge since there are only 3 KA crashes sampled. Also, in Ohio's SV crashes' SPF, the constant is positive and the natural logarithm of the AADT is insignificant at the 95th percentile confidence interval. Likewise, that may be because of the low SV crash frequency. There are only 17 SV crashes sampled from Ohio's four-leg signalized intersections conforming to the proposed base conditions.

7.2.2 Florida and Ohio's Joint Proposed Base Conditions Rural Four-Leg Signalized Intersections' Safety Performance Functions

The joint Florida and Ohio SPFs are estimated for KABCO, KABC and MV crashes. The models for KAB, KA and SV crashes failed to converge possibly because of the low crash frequencies. The joint SPF results are shown in Table 7.11.

Table 7.11: Florida and Ohio's Joint SPFs for Proposed Base Conditions Rural Four-Leg Signalized Intersections

Crash Type or Severity	**KABCO**	**KABC**	**KAB**	**KA**	**SV**	**MV**
Parameters	**Parameter Estimates and P-Values**					
Constant	-7.3298 (<.0001)	-9.5407 (0.0012)	Failed to converge	Failed to converge	Failed to converge	-8.4582 (<.0001)
Ln(AADT major)	0.5352 (0.0002)	NA				0.6575 (<.0001)
Ln(AADT minor)	0.1804 (0.0077)	NA				0.1735 (0.0119)
Ln(AADT total)	NA	0.9264 (0.0035)				NA
OH	1.9745 (<.0001)	-				1.9934 (<.0001)
OH ×Ln (AADT major)	-	NA				-
OH ×Ln (AADT minor)	-	NA				-
OH ×Ln (AADT total)	NA	-				NA
P	2.2116 (0.0660)	1.8256 (0.2727)				1.9050 (0.0835)
Q	-2.0035 (0.0962)	-1.0740 (0.5523)				-1.6885 (0.1258)
Goodness of Fit Measures						
-2LL	375.6	235.8	Failed to converge	Failed to converge	Failed to converge	365.5
MAD	1.805	0.839				1.729
MSPE	6.696	1.819				6.450
	NA : not applicable variables - : statistically insignificant variables at alpha = 0.05 removed from the SPF					

As shown in Table 7.11, the interaction terms between the dummy variable representing Ohio, *OH*, and the natural logarithm of the entering AADTs are statistically insignificant at the 95 percentile confidence interval for all crash classifications. It is an indication that for all crash types and severity levels, the impact of the entering AADTs on crashes in Ohio is the same as that in Florida. It should also be noted that the insignificance of the dummy variable, *OH*, in the KABC crashes' SPF indicates that for the same entering AADTs, in both states, Ohio's intersections experience the same KABC crash counts as Florida's. That is a consistent result with the KABC SPF of the joint Florida and Ohio average conditions four-leg signalized intersections data. For KABCO and MV crashes' SPFs the dummy variable is statistically significant implying that the intersections in Ohio experience more KABCO and MV crashes relative to those of Florida given the same entering AADT.

7.2.3 Transferability Assessment for Rural Four-Leg Signalized Intersections' Proposed Base Conditions

The Florida SPFs are applied to Ohio and vice versa. Also, the joint SPFs are applied to both states and the transfer indices are calculated accordingly. The transfer index results are shown in Table 7.12.

Table 7.12: Proposed Base Conditions Rural Four-Leg Signalized Intersections - Transfer Indices

		Application Data	
Crash Type or Severity	**SPF**	**Florida**	**Ohio**
KABCO	**Florida**	1	-16.381
	Ohio	-25.964	1
	Florida and Ohio	-1.033	0.994
KABC	**Florida**	NA	NA
	Ohio	NA	1
	Florida and Ohio	NA	0.755
KAB	**Florida**	NA	NA
	Ohio	NA	NA
	Florida and Ohio	NA	NA
KA	**Florida**	NA	NA
	Ohio	NA	NA
	Florida and Ohio	NA	NA
SV	**Florida**	1	NA
	Ohio	-160.880	1
	Florida and Ohio	NA	NA
MV	**Florida**	1	-14.973
	Ohio	-18.918	1
	Florida and Ohio	-0.318	0.993
	NA : Transfer index not calculated because of corresponding SPF failure to converge		

The transfer indices are negative indicating that the Florida four-leg signalized intersection SPFs are not transferable to Ohio and vice versa. That is a consistent result with the case of four-leg signalized intersections conforming to the average conditions.

7.2.4 Application of the Modified Empirical Bayes Method to Proposed Base Conditions Rural Four-Leg Signalized Intersections

The modified EB method is applied for correcting crash predictions of SPFs of four-leg signalized intersections conforming to the base conditions. All developed SPFs are applied to Florida and Ohio jurisdictions to obtain the predicted crash frequencies. Then, the average of the observed crash frequencies of the years 2009 and 2010 are calculated for every intersection. That is to be able to apply the average observed crash frequencies to obtain the expected ones. Also, the HSM calibration factors are obtained using the average observed crash frequencies and the predicted ones. Then, the MAD and MSPE measures are calculated by comparing the predicted crash frequencies with the observed ones of the validation data. The measures are also calculated to compare the predicted crash frequencies adjusted by the HSM calibration factors with the observed crash frequencies of the validation data. Finally, the comparison is made between the expected and observed crash frequencies of the validation data. The comparison results are shown in Tables 7.13 and 7.14.

Table 7.13: Comparison of Predicted, Calibrated and Expected Crash Frequencies in Florida – Proposed Base Conditions Rural Four-Leg Signalized Intersections

Crash Type or Severity	SPF	Application Data: Florida 2011					
		MAD	MSPE	MAD Calculated Using HSM Calibration Factor	MSPE Calculated Using HSM Calibration Factor	MAD Calculated Using Modified EB	MSPE Calculated Using Modified EB
KABCO	Florida	2.221	15.531	2.218	15.487	2.19	15.029
	Ohio	2.404	9.004	2.303	16.611	2.175	10.545
	Florida and Ohio	2.309	16.765	2.298	16.543	2.236	15.532
KABC	Florida	NA	NA	NA	NA	NA	NA
	Ohio	0.95	2.425	0.969	3.812	0.965	2.941
	Florida and Ohio	0.988	3.037	0.979	3.872	0.953	3.361
KAB	Florida	NA	NA	NA	NA	NA	NA
	Ohio	0.488	0.407	0.483	0.855	0.501	0.58
	Florida and Ohio	NA	NA	NA	NA	NA	NA
KA	Florida	NA	NA	NA	NA	NA	NA
	Ohio	NA	NA	NA	NA	NA	NA
	Florida and Ohio	NA	NA	NA	NA	NA	NA
SV	Florida	0.139	0.127	0.139	0.127	0.139	0.127
	Ohio	0.211	0.115	0.138	0.125	0.203	0.116
	Florida and Ohio	NA	NA	NA	NA	NA	NA
MV	Florida	2.096	14.352	2.096	14.349	2.065	14.078
	Ohio	2.324	8.186	2.2	15.582	2.11	9.867
	Florida and Ohio	2.203	15.734	2.193	15.499	2.121	14.646
	NA : Measure not calculated because of corresponding SPF failure to converge						

Table 7.14: Comparison of Predicted, Calibrated and Expected Crash Frequencies in Ohio – Proposed Base Conditions Rural Four-Leg Signalized Intersections

Crash Type or Severity	SPF	Application Data: Ohio 2011					
		MAD	MSPE	MAD Calculated Using HSM Calibration Factor	MSPE Calculated Using HSM Calibration Factor	MAD Calculated Using Modified EB	MSPE Calculated Using Modified EB
KABCO	Florida	2.245	8.673	1.498	4.494	1.574	4.222
	Ohio	1.286	3.177	1.29	3.133	1.019	2.022
	Florida and Ohio	1.287	3.169	1.289	3.146	1.016	1.987
KABC	Florida	NA	NA	NA	NA	NA	NA
	Ohio	0.779	1.295	0.783	1.265	0.699	1.081
	Florida and Ohio	0.785	1.459	0.788	1.291	0.716	1.193
KAB	Florida	NA	NA	NA	NA	NA	NA
	Ohio	0.606	0.862	0.61	0.854	0.543	0.701
	Florida and Ohio	NA	NA	NA	NA	NA	NA
KA	Florida	NA	NA	NA	NA	NA	NA
	Ohio	NA	NA	NA	NA	NA	NA
	Florida and Ohio	NA	NA	NA	NA	NA	NA
SV	Florida	0.137	0.107	0.239	0.251	0.137	0.107
	Ohio	0.211	0.102	0.25	0.106	0.212	0.101
	Florida and Ohio	NA	NA	NA	NA	NA	NA
MV	Florida	2.155	8.17	1.516	4.723	1.625	4.479
	Ohio	1.246	3.092	1.245	3.08	0.999	1.918
	Florida and Ohio	1.245	3.102	1.245	3.108	0.99	1.876
	NA : Measure not calculated because of corresponding SPF failure to converge						

When Florida's validation data are applied, the results of the modified EB method are superior to those of the HSM calibration method except for KAB, KA and SV crashes. The case is not the same when Ohio's validation data are applied. Instead, the MADs obtained using the modified EB method are higher than those obtained from the HSM calibration method only for Florida's KABCO crashes' SPF and Florida's MV crashes' SPF.

CHAPTER 8: CONCLUSIONS

The SPFs are indispensable for enhancing traffic safety by enabling the prediction of crash frequencies and screen road networks for hot spots. The SPFs may be applied to predict crash frequencies for any road user, whether vehicle driver, pedestrian or bicyclist. Also, the predicted crashes may be of any type or severity level whether fatal crashes K, incapacitating injury crashes, A, non-incapacitating injury crashes, B, possible injury crashes, C, or property damage only crashes, O. In the HSM, several NB SPFs are provided for different types of roadway facilities. The HSM SPFs are applied particularly for regions where jurisdiction specific SPFs are not developed. The HSM SPFs, applied to the region of interest, are modified by calibration factors. The factors are calculated based on the ratio of the sum of the observed crashes to that of the predicted crashes for every segment or intersection. This approach has been active lately in traffic safety research. However, there has been limited research in the exploration of the soundness of SPF transferability.

Considering an intricate assessment of transferability of SPFs among multiple regions contributes to the evolving literature of SPF transferability. Specifically, in this study, jurisdiction specific SPFs of Florida, Ohio and California are applied to each state and the prediction accuracies of the SPFs are compared to fulfill the objective of assessing the transferability of SPFs among the three states. Jurisdiction specific SPFs are estimated for KABCO, KABC, KAB, KA, SV and MV crashes. Then, each state's data are pooled in different combinations to develop two-state and three-state joint SPFs for all crash classifications. That is performed for rural divided multilane highway segments and four-leg signalized intersections. The analysis is undertaken once more for proposed base conditions. That allows for controlling for roadway characteristics such as the lane width, shoulder width, median width and lighting conditions for divided segments. For the four-leg signalized intersections, the controlled variables are the skew angle and street lighting presence only. These are not the defined base conditions of the HSM since there is an inadequate sample of divided segments that satisfy the HSM base conditions. Also, there are currently no base conditions defined in the HSM for four-leg signalized intersections. There are two other objectives in this study. One is to investigate the effect of pooling data from multiple states and estimating joint SPFs on the SPF transferability. The results of the joint SPFs are compared with those estimated from data of single states. The transfer index, a measure that was previously applied in transportation research is used to assess transferability. The superior transfer indices of the two-state and three-state joint SPFs relative to those of the one-state SPFs clearly indicate that pooling data to develop SPFs improves SPF transferability. Among the three states under study, Florida's SPFs for divided segments conforming to the average conditions are more transferable to California and vice versa compared to those of Ohio's. The final objective is to explore the applicability of the proposed disaggregate adjustment procedure that corrects predictions of SPFs being transferred to the region of interest. It is the modified EB method. Results of the comparison between the modified EB method with the HSM calibration method demonstrate the validity of the proposed method. Also according to the results, the disaggregate EB method outperforms the HSM aggregate calibration method especially for segments. For four-leg signalized intersections, the EB method performs better than the HSM calibration method in most cases possibly because the overdispersion parameter is fixed for every intersection.

8.1 Key Findings

The key findings of this study are summarized as follows.

- The transfer indices obtained using joint SPFs are higher than those obtained from jurisdiction specific SPFs. As previously stated in the introduction chapter, the HSM's SPFs for rural four-leg signalized intersections are developed from data of Minnesota while the SPFs of rural divided multilane highway segments are developed from pooled data of California and Texas to satisfy a reasonable sample size. Even though the sample size of Minnesota's four-leg signalized intersections is sufficient it is recommended to pool the data with those of another state. Likewise, the Texas and California rural divided segments data can be pooled with a third state before estimating the SPFs. The resulting joint SPFs capture characteristics of multiple states and therefore become more transferable to a particular jurisdiction of interest especially if the jurisdiction is near states of which data are used for estimation of the joint SPFs.
- Florida's SPFs of rural divided segments conforming to average conditions are not transferable to Ohio and vice versa. On the other hand, the case is opposite for Florida and California's SPFs except for SV crashes.
- Ohio's average conditions divided segments KABCO, SV and MV crashes' SPFs are transferable to California. Similarly, those of California are transferable to Ohio.
- Under the modified base conditions, Ohio's KABC and KAB crashes' SPFs are transferable to California and vice versa. The corresponding transfer indices are higher than their counterparts for divided segments conforming to average conditions.

- Adding coefficients for the dummy variable representing the state in the overdispersion parameter formula accounts for differences in crash patterns among states when estimating joint SPFs.
- According to the transfer index results of both average and proposed base conditions of four-leg signalized intersections, neither jurisdiction specific SPFs are transferable to any state for all crash classifications. That is also true for the joint SPFs.
- Even though Florida's rural divided segments' SPFs are transferable to California and California's are transferable to Florida, both states are different in many dimensions. Both states are not only different in terms of geographic locations but also in terms of topography and weather conditions. Florida represents the southeast with its floundering vegetation and intense rainfall while California represents the southwest with its mountainous terrain and considerably less rainfall. However the majority of California's population do not reside in mountainous territories. When it comes to similarities in both states, the inhabited areas in California do not experience snow as Florida does not either. Also, the roadway facilities at mountainous areas in California are not included in the data for this study. In addition, tourism is common in both states. On the other hand, Florida and Ohio are different especially in terms of weather conditions since Ohio experiences snow during Winter while Florida experiences intense rainfall in Summer and Fall. Crash reporting thresholds are different in all three states. In terms of property damage, the thresholds are $500 in Florida as per the Florida Statutes section governing traffic, $400 in Ohio as per the Ohio Bureau of Motor Vehicles and $750 in California (Xie *et al*., 2011). Yet, the minimum property damage, caused by traffic crashes, for the crashes to be reported in Ohio was changed to $1000 in 2011 as per the Ohio data

description in the HSIS. According to the Transportation Motor Vehicle Accidents and Fatalities Summaries of US Census Bureau, the national crash rate for 2009 is 330.9 total crashes per hundred million VMT. The national rate of fatal crashes for the same year is 1.04 fatal crashes per hundred million VMT according to the Fatality Analysis Reporting System of the National Highway Traffic Safety Administration (NHTSA) Encyclopedia. As per the Florida Integrated Report Exchange System of the Department of Highway Safety and Motor Vehicles, the rates are 119.59 total crashes per hundred million VMT and 1.20 fatal crashes per hundred million VMT for 2009. Those of Ohio for the same year are 269.61 total crashes per hundred million VMT and 0.85 fatal crashes per hundred million VMT as per the Access Ohio 2040 program of the Ohio Department of Transportation. In California, the rate is 131.42 total crashes per hundred million VMT in 2009 as per the National Center for Statistics and Analysis Data Resource. According to the General Statistics of the Insurance Institute for Highway Safety there are 2,816 fatal crashes reported in California the same year while the vehicle mileage traveled is 3,243 hundred million VMT as per the California Traffic Safety Score Card. That is 0.87 fatal crashes per hundred million VMT.

8.2 Future Work

Unquestionably, there are limitations to this study. First, according to the key findings, the modified EB method, proposed, outperforms the HSM calibration method in the majority of cases but there is potential for improvement. That is, further research is required to enhance the performance of the proposed EB method. Also, with the availability of data of previous crash years, research may be conducted to examine the effectiveness of constructing Bayesian informative priors to expand upon previous studies' work such as that of Yu and Abdel-Aty (2013). In addition, it is recommended to examine the effect of applying joint SPFs, developed from three states, into data of a fourth state. That is because transfer indices corresponding to joint SPFs are greater than those that belong to SPFs developed from a single state according to the key findings of this study. In addition, the NB SPFs applied are of limited variables since only the AADT is considered as a variable. Variables not considered such as lane width, shoulder width, median width, grades, intersection skew angle, horizontal curvature, posted speed limits, roadside hazard rating, other geometric characteristics, alcohol or drug use, weather conditions or any other factors are not included in the SPFs. Including such variables will improve the estimated SPFs and most likely facilitate transferability. It would be convenient to extend the work carried out for rural divided multilane highway segments and four-leg signalized intersections to all other types of roadway facilities. Finally, in this study, the transferability of SPFs is conducted for distant states representing different regions. It would be recommended to re-conduct the assessment for neighbor states within the same region. Specifically, future research is suggested for the transferability assessment of jurisdiction specific SPFs to neighbor states in regions with similar demographic patterns, topography and weather trends. For instance, the transferability of SPFs of Ohio and Indiana may be examined since both states are adjacent to each other.

LIST OF REFERENCES

Abdel-Aty, M., and J. Keller. Exploring the Overall and Specific Crash Severity Levels at Signalized Intersections. *Accident Analysis & Prevention*, Vol. 37, No. 3, 2005, pp. 417-425.

Access Ohio 2040. *Technical Memorandum: Safety*. Ohio Department of Transportation, Columbus, Ohio. http://www.dot.state.oh.us/Divisions/Planning/SPR/StatewidePlanning/access.ohio/Documents/TechMemos/Safety.pdf. Accessed Sept. 30, 2015.

Ackaah, W., and M. Salifu. Crash Prediction Model for Two-Lane Rural Highways in the Ashanti Region of Ghana. *IATSS Research*, Vol. 35, No. 1, 2011, pp. 34-40.

Al Kaaf, K., and M. Abdel-Aty. Transferability and Calibration of Highway Safety Manual Performance Functions and Development of New Models for Urban Four-Lane Divided Roads in Riyadh. Presented at 94th Annual Meeting of the Transportation Research Board, Washington, D.C., 2015.

American Association of State Highway and Transportation Officials. *Highway Safety Manual*, Washington, DC, 2010.

Atherton, T. and M. Ben-Akiva. Transferability and Updating of Disaggregate Travel Demand Models. In *Transportation Research Record: Journal of the Transportation Research Board, No. 610*, Transportation Research Board of the National Academies, Washington, D.C., 1976, pp. 12-18.

Cafiso, S., G. Di Silvestro, and G. Di Guardo. Application of Highway Safety Manual to Italian Divided Multilane Highways. *Procedia-Social and Behavioral Sciences*, Vol. 53, 2012, pp. 910-919.

California Traffic Safety Score Card. *Overall*. California Office of Traffic Safety, Elk Grove, California. http://www.ots.ca.gov/OTS_and_Traffic_Safety/Score_Card_2011.asp. Accessed Sept. 30, 2015.

Donnell, E., and J. Mason. Predicting the Frequency of Median Barrier Crashes on Pennsylvania Interstate Highways. *Accident Analysis & Prevention*, Vol. 38, No. 3, 2006, pp. 590-599.

Driver Safety and Laws. *Crash Reports and Records*. Texas Department of Transportation, Austin, Texas. https://www.google.com/webhp?sourceid=chrome-instant&ion=1&espv=2&ie=UTF-8#q=Txdot+hq. Accessed Oct. 27, 2015.

Farid, A., M. Abdel-Aty, J. Lee, N. Eluru, and J.-H. Wang. Exploring Transferability of Safety Performance Functions: A Case Study of Rural Multilane Divided Highway Segments. Accepted for Presentation at 95th Annual Meeting of the Transportation Research Board, Washington, D.C., 2016.

Fatality Analysis Reporting System of the National Highway Traffic Safety Administration Encyclopedia. *National Statistics*. National Highway Traffic Safety Administration, United States Department of Transportation, Washington, D.C. http://www-fars.nhtsa.dot.gov/Main/index.aspx. Accessed Sept. 30, 2015.

Florida Integrated Report Exchange System. *Traffic Crash Statistics Report 2009*. Florida Department of Highway Safety and Motor Vehicles, Tallahassee, Florida.

https://firesportal.com/Pages/Public/DHSMVPublishedDocuments/Previous%20Years/Crash%20Facts%202009.pdf. Accessed Sept. 30, 2015.

Florida Statutes. *Title XXIII – Motor Vehicles*. The Florida Legislature, Tallahassee, Florida. http://www.leg.state.fl.us/Statutes/index.cfm?App_mode=Display_Statute&Search_String=&URL=0300-0399/0316/Sections/0316.065.html. Accessed Oct. 28, 2015.

Garber, N., and L. Wu. Stochastic Models Relating Crash Probabilities with Geometric and Corresponding Traffic Characteristics Data. Presented to National Intelligent Transportation Systems Implementation Research Center of the Department of Transportation, Charlottesville, 2001.

General Statistics. *Population, Fatal Motor Vehicle Crashes, Motor Vehicle Crash Deaths, and Motor Vehicle Crash Deaths per 100,000 People by State, 2009*. Insurance Institute for Highway Safety - Highway Loss Data Institute, Arlington, Virginia. http://www.iihs.org/iihs/topics/t/general-statistics/fatalityfacts/state-by-state-overview/2009#Crash-types. Accessed Sept. 30, 2015.

Hadayeghi, A., A. Shalaby, B. Persaud, and C. Cheung. Temporal Transferability and Updating of Zonal Level Accident Prediction Models. *Accident Analysis & Prevention*, Vol. 38, No. 3, 2006, pp. 579-589.

Harnen, S., R. Umar, S. Wong, and W. Hashim. Development of Prediction Models for Motorcycle Crashes at Signalized Intersections on Urban Roads in Malaysia. *Journal of Transportation and Statistics*, Vol. 7, No. 2, 2004, p. 27.

Hauer, E., and J. Bamfo. Two Tools for Finding What Function links the dependent Variable to the Explanatory Variables. Presented at 10th Annual Meeting of the International Co-operation on Theories and Concepts in Traffic Safety, Vienna, Austria, 1997.

Highway Safety Information System. *Guidebook for the Ohio State Data Files*. Highway Safety Research Center, University of North Carolina, Chapel Hill, North Carolina. http://www.hsisinfo.org/guidebooks/ohio.cfm. Accessed Nov. 11, 2015.

Highway Safety Information System. *State Data*. Highway Safety Research Center, University of North Carolina, Chapel Hill, North Carolina. http://www.hsisinfo.org/. Accessed Sept. 28, 2015.

Huang, H., C. Siddiqui, and M. Abdel-Aty. Indexing Crash Worthiness and Crash Aggressivity by Vehicle Type. *Accident Analysis & Prevention*, Vol. 43, No. 4, 2011, pp. 1364-1370.

Jovanis, P., and F. Gross. Estimation of Safety Effectiveness of Changes in Shoulder Width with Case Control and Cohort Methods. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2019*, Transportation Research Board of the National Academies, Washington, D.C., 2007, pp. 237-245.

Kim, S., S. Chung, K. Song, and K. Chon. Development of an Accident Prediction Model using GLIM (Generalized Log-linear Model) and EB method: A case of Seoul. *Journal of the Eastern Asia Society for Transportation Studies*, Vol. 6, 2005, pp. 3669-3682.

Local and Rural Road Safety Program. *2012 Traffic Safety Facts "Rural and Urban Comparison" DOT-HS-812-050*. Federal Highway Administration, United States Department of Transportation, Washington D.C. http://safety.fhwa.dot.gov/local_rural/#programs. Accessed July 7, 2015.

Lord, D., S. Geedipally, B. Persaud, S. Washington, I. van Schalkwyk, J. Ivan, C. Lyon, and T. Jonsson. *NCHRP 126 Report: Methodology for Estimating the Safety Performance of Rural Multilane Highways*. National Cooperative Highway Research Program, Transportation Research Board of the National Academies, Washington, D.C., 2008.

Lord, D., S. Washington, and J. Ivan. Poisson, Poisson-gamma and Zero-Inflated Regression Models of Motor Vehicle Crashes: Balancing Statistical Fit and Theory. *Accident Analysis & Prevention*, Vol. 37, No. 1, 2005, pp. 35-46.

Mehta, G., and Y. Lou. Calibration and Development of Safety Performance Functions for Alabama: Two-Lane, Two-Way Rural Roads and Four-Lane Divided Highways. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2398*, Transportation Research Board of the National Academies, Washington, D.C., 2013, pp. 75-82.

Miaou, S.-P. The Relationship between Truck Accidents and Geometric Design of Road Sections: Poisson versus Negative Binomial Regressions. *Accident Analysis & Prevention*, Vol. 26, No. 4, 1994, pp. 471-482.

Miaou, S.-P., and H. Lum. Modeling Vehicle Accidents and Highway Geometric Design Relationships. *Accident Analysis & Prevention*, Vol. 25, No. 6, 1993, pp. 689-709.

National Center for Statistics and Analysis Data Resource Website: Customer Automated Tracking System. *State Data System Crash Data Report: 2000-2009*. National Highway Traffic Safety Administration, United States Department of Transportation, Washington, D.C. http://www-nrd.nhtsa.dot.gov/Pubs/812052.pdf. Accessed Sept. 30, 2015.

Ohio Bureau of Motor Vehicles. *When Involved in an Automobile Crash*. Ohio Department of Public Safety, Columbus, Ohio. http://www.bmv.ohio.gov/fr_laws.stm. Accessed Oct. 28, 2015.

Pande, A., and M. Abdel-Aty. A Novel Approach for Analyzing Severe Crash Patterns on Multilane Highways. *Accident Analysis & Prevention*, Vol. 41, No. 5, 2009, pp. 985-994.

Persaud, B., D. Lord, and J. Palmisano. Calibration and Transferability of Accident Prediction Models for Urban Intersections. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1784*, Transportation Research Board of the National Academies, Washington, D.C., 2002, pp. 57–64.

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

Salifu, M. Accident Prediction Models for Unsignalised Urban Junctions in Ghana. *IATSS Research*, Vol. 28, No. 1, 2004, pp. 68-81.

Sawalha, Z., and T. Sayed. Traffic Accident Modeling: Some Statistical Issues. *Canadian Journal of Civil Engineering*, Vol. 33, No. 9, 2006, pp. 1115-1124.

Sikder, S., B. Augustin, A. Pinjari, and N. Eluru. Spatial Transferability of Tour-Based Time-of-Day Choice Models: An Empirical Assessment. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2429*, Transportation Research Board of the National Academies, Washington, D.C., 2014, pp. 99–109.

Srinivasan, R., and D. Carter. *Development of Safety Performance Functions for North Carolina*. Publication FHWA/NC/2010-09. Research and Analysis Group, North Carolina Department of Transportation, 2011.

Sun, C., H. Brown, P. Edara, B. Claros, and K. Nam. *Calibration of the HSM's SPFs for Missouri*. Publication CMR14-007. Missouri Department of Transportation, 2014.

Taylor, M., A. Baruya, and J. Kennedy. *The Relationship between Speed and Accidents on Rural Single-Carriageway Roads*. Publication TRL511. Transport Research Laboratory, 2002.

Transportation: Motor Vehicle Accidents and Fatalities. *Motor Vehicle Accidents – Number and Deaths*. United States Census Bureau, Suitland, Maryland. http://www.census.gov/compendia/statab/2012/tables/12s1103.pdf. Accessed Sept. 30, 2015.

Turner, S., B. Persaud, and M. Chou. Transferability of Overseas Crash Prediction Models to New Zealand. Presented at 11th Annual Meeting of the Transportation Group Conference of the Institution of Professional Engineers New Zealand, Auckland, N.Z., 2007.

Xie, F., K. Gladhill, K. Dixon, and C. Monsere. Calibration of Highway Safety Manual Predictive Models for Oregon State Highways. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2241*, Transportation Research Board of the National Academies, Washington, D.C., 2011, pp. 19-28.

Yan, X., M. Abdel-Aty, E. Radwan, X. Wang, and P. Chilakapati. Validating a Driving Simulator Using Surrogate Safety Measures. *Accident Analysis & Prevention*, Vol. 40, No. 1, 2008, pp. 274-288.

Young, J., and P. Park. Comparing the Highway Safety Manual's Safety Performance Functions with Jurisdiction-Specific Functions for Intersections in Regina. Presented at 10th Annual Meeting of the Transportation Association of Canada, Ottawa, Ontario, Canada, 2012.

Yu, R., and M. Abdel-Aty. Investigating Different Approaches to Develop Informative Priors in Hierarchical Bayesian Safety Performance Functions. *Accident Analysis & Prevention*, Vol. 56, 2013, pp. 51-58.