Developing Crash Modification Factors for Bicycle-Lane Additions While Reducing Lane and Shoulder Widths

PUBLICATION NO. FHWA-HRT-21-013 JULY 2021

U.S. Department of Transportation
Federal Highway Administration

Research, Development, and Technology
Turner-Fairbank Highway Research Center
6300 Georgetown Pike
McLean, VA 22101-2296

FOREWORD

The research documented in this report was conducted as part of the Federal Highway Administration's (FHWA's) Evaluation of Low-Cost Safety Improvements Pooled Fund Study (ELCSI-PFS). FHWA established this PFS in 2005 to conduct research on the effectiveness of safety improvements identified by the National Cooperative Highway Research Program's Report 500 Series as part of the implementation of the American Association of State Highway and Transportation Officials' *Strategic Highway Safety Plan* (AASHTO 1998; TRB 2003). The ELCSI-PFS provides a crash modification factor (CMF) and benefit–cost (B/C) economic analysis for each of the targeted safety strategies identified as priorities by the PFS member States.

This study evaluated the safety effectiveness of bicycle-lane additions by reducing lane and shoulder width as a safety-improvement strategy. CMFs and B/C ratios were estimated for total, fatal and injury, and property-damage-only (PDO) crashes. The study included urban two- and four-lane arterials, collectors, and local roads in Washington and Texas. Results for Washington suggested reductions for fatal and injury and PDO crashes but were statistically insignificant. Results for Texas were consistent with the Washington analysis and showed a statistically significant reduction in total and PDO crashes. The B/C ratio of 16.61 was estimated only for Texas two-lane undivided urban collectors and local streets because of a larger sample size.

The results of this study may be of interest to State and local engineers and planners responsible for the design, operation, and maintenance of bicycle facilities. Highway-safety practitioners can use the results to make decisions regarding the safety effectiveness of adding bicycle lanes while reducing lane and shoulder widths.

Brian P. Cronin, P.E.
Director, Office of Safety and Operations
Research and Development

TECHNICAL REPORT DOCUMENTATION PAGE

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
FHWA-HRT-21-013		

4. Title and Subtitle
Developing Crash Modification Factors for Bicycle-Lane Additions While Reducing Lane and Shoulder Widths

5. Report Date
July 2021

6. Performing Organization Code

7. Author(s)
Raul Avelar (ORCID: 0000-0002-3962-1758), Karen Dixon (ORCID: 0000-0002-8431-9304), Sruthi Ashraf (ORCID: 0000-0002-3304-9682), Ankit Jhamb (ORCID: 0000-0002-6123-9352), and Bahar Dadashova (ORCID: 0000-0002-4592-9118)

8. Performing Organization Report No.

9. Performing Organization Name and Address
Texas A&M Transportation Institute
The Texas A&M University System
College Station, TX 77843-3135

10. Work Unit No.

11. Contract or Grant No.
DTFH6116D00039

12. Sponsoring Agency Name and Address
Office of Safety Research and Development
Federal Highway Administration
6300 Georgetown Pike
McLean, VA 22101-2296

13. Type of Report and Period Covered
Final Report; May 2017–April 2021

14. Sponsoring Agency Code
HRDS-20

15. Supplementary Notes
The Federal Highway Administration Development of Crash Modification Factors Program and Task Manager for this project was Roya Amjadi (HRDS-20; ORCID: 0000-0002-8426-0867).

16. Abstract
This project evaluated bicycle-lane additions when reducing lane and shoulder width as a safety-improvement strategy (also known as a safety intervention). Crash modification factors (CMFs) and benefit–cost (B/C) ratios were estimated for total, fatal and injury, and property-damage-only (PDO) crashes. Researchers studied facilities that included urban two- and four-lane arterials, collectors, and local roads in Washington and Texas. The study design was cross sectional and included sites without bicycle lanes but with similar characteristics to sites with bicycle lanes. The research team applied propensity score methods on data collection and in statistical analyses to improve the balance between treated and comparison sites. Results for Washington suggested reductions for fatal and injury (0.772 CMF, statistically insignificant) and PDO (0.885 CMF, statistically insignificant) crashes, but the small sample size resulted in large uncertainties for these estimates.

Results for Texas were consistent with the directions and magnitudes of the Washington analysis. In contrast, the larger dataset from Texas yielded CMF estimates with statistical significance at the 95-percent confidence level. From the Texas analysis, this research found that statistically significant CMFs for total and PDO crashes ranged between 0.514 and 0.734, depending on facility type and bicycle volumes. All CMFs for fatal and injury crashes were statistically insignificant (ranging between 0.530 and 1.58, depending on facility type and bicycle volumes), as were CMFs for total crashes (ranging from 0.558 to 0.901, all statistically insignificant). The B/C ratio estimated for Texas two-lane undivided urban collectors and local streets was 16.61, assuming the average daily bicycle traffic did not increase after the installation of bicycle lanes.

17. Key Words
Safety, crash modification factors, CMFs, bicycle lanes

18. Distribution Statement
No restrictions. This document is available to the public through the National Technical Information Service, Springfield, VA 22161.
http://www.ntis.gov

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	61	N/A

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized.

SI* (MODERN METRIC) CONVERSION FACTORS				
APPROXIMATE CONVERSIONS TO SI UNITS				
Symbol	**When You Know**	**Multiply By**	**To Find**	**Symbol**
		LENGTH		
in	inches	25.4	millimeters	mm
ft	feet	0.305	meters	m
yd	yards	0.914	meters	m
mi	miles	1.61	kilometers	km
		AREA		
in^2	square inches	645.2	square millimeters	mm^2
ft^2	square feet	0.093	square meters	m^2
yd^2	square yard	0.836	square meters	m^2
ac	acres	0.405	hectares	ha
mi^2	square miles	2.59	square kilometers	km^2
		VOLUME		
fl oz	fluid ounces	29.57	milliliters	mL
gal	gallons	3.785	liters	L
ft^3	cubic feet	0.028	cubic meters	m^3
yd^3	cubic yards	0.765	cubic meters	m^3
		NOTE: volumes greater than 1,000 L shall be shown in m^3		
		MASS		
oz	ounces	28.35	grams	g
lb	pounds	0.454	kilograms	kg
T	short tons (2,000 lb)	0.907	megagrams (or "metric ton")	Mg (or "t")
		TEMPERATURE (exact degrees)		
°F	Fahrenheit	5 (F-32)/9 or (F-32)/1.8	Celsius	°C
		ILLUMINATION		
fc	foot-candles	10.76	lux	lx
fl	foot-Lamberts	3.426	candela/m^2	cd/m^2
		FORCE and PRESSURE or STRESS		
lbf	poundforce	4.45	newtons	N
lbf/in^2	poundforce per square inch	6.89	kilopascals	kPa
APPROXIMATE CONVERSIONS FROM SI UNITS				
Symbol	**When You Know**	**Multiply By**	**To Find**	**Symbol**
		LENGTH		
mm	millimeters	0.039	inches	in
m	meters	3.28	feet	ft
m	meters	1.09	yards	yd
km	kilometers	0.621	miles	mi
		AREA		
mm^2	square millimeters	0.0016	square inches	in^2
m^2	square meters	10.764	square feet	ft^2
m^2	square meters	1.195	square yards	yd^2
ha	hectares	2.47	acres	ac
km^2	square kilometers	0.386	square miles	mi^2
		VOLUME		
mL	milliliters	0.034	fluid ounces	fl oz
L	liters	0.264	gallons	gal
m^3	cubic meters	35.314	cubic feet	ft^3
m^3	cubic meters	1.307	cubic yards	yd^3
		MASS		
g	grams	0.035	ounces	oz
kg	kilograms	2.202	pounds	lb
Mg (or "t")	megagrams (or "metric ton")	1.103	short tons (2,000 lb)	T
		TEMPERATURE (exact degrees)		
°C	Celsius	1.8C+32	Fahrenheit	°F
		ILLUMINATION		
lx	lux	0.0929	foot-candles	fc
cd/m^2	candela/m2	0.2919	foot-Lamberts	fl
		FORCE and PRESSURE or STRESS		
N	newtons	2.225	poundforce	lbf
kPa	kilopascals	0.145	poundforce per square inch	lbf/in^2

*SI is the symbol for International System of Units. Appropriate rounding should be made to comply with Section 4 of ASTM E380. (Revised March 2003)

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS AND ABBREVIATIONS

AADT	average annual daily traffic
AASHTO	American Association of State Highway and Transportation Officials
ADBT	average daily bicycle traffic
B/C	benefit–cost
CMF	crash modification factor
DCMF	Development of Crash Modification Factors
DOT	department of transportation
EB	empirical Bayes
ELCSI-PFS	Evaluation of Low-Cost Safety Improvements Pooled Fund Study
FHWA	Federal Highway Administration
GIS	geographic information systems
GLM	generalized linear model
GLMM	generalized linear mixed model
GNM	generalized nonlinear model
KABC	fatal and injury crashes (part of the KABCO scale)
NACTO	National Association of City Transportation Officials
NBPDP	National Bicycle and Pedestrian Documentation Project
PDO	property damage only
PEM	passing event model
PS	propensity score
PSM	propensity score matching
PSW	propensity score weighting
RID	Roadway Inventory Database
SE	standard error
SPF	safety performance function
USDOT	U.S. Department of Transportation
VSL	value of a statistical life
WCL	wide curb lane

EXECUTIVE SUMMARY

The Federal Highway Administration's (FHWA's) Development of Crash Modification Factors (DCMF) program was established in 2012 to address highway-safety research needs for evaluating new and innovative safety strategies (e.g., improvements) by developing reliable quantitative estimates of their effectiveness in reducing crashes.

The goal of FHWA's DCMF program is to save lives by identifying new safety strategies that effectively reduce crashes and promoting the strategies for nationwide installation by providing measures of their safety effectiveness and benefit–cost (B/C) ratios through research. State departments of transportation (DOTs) and other transportation agencies must have objective measures for safety effectiveness and B/C ratios before investing in new strategies for statewide safety improvements.

Forty-one State DOTs provided technical feedback on safety improvements to the DCMF program and implemented new safety improvements to facilitate evaluations. These States are members of the Evaluation of Low-Cost Safety Improvements Pooled Fund Study (ELCSI-PFS) that functions under the DCMF program.

This project evaluated bicycle-lane additions accomplished by reducing lane and shoulder width as a safety-improvement strategy (safety intervention). The ELCSI-PFS Technical Advisory Committee selected the evaluation of bicycle-lane additions accomplished by reducing lane and shoulder width as one of the priorities within its purview.

This evaluation assessed the potential of the safety-improvement strategy (i.e., adding bicycle lanes while reducing shoulder and lane widths) to reduce crashes in terms of total, fatal and injury, property-damage-only (PDO), and bicycle crash frequencies. The intent is to develop crash modification factors (CMFs) and B/C ratios for the safety-improvement strategy. Practitioners can use CMFs and B/C ratios when deciding on the project development and safety-planning processes.

This study evaluated bicycle-lane additions on urban two- and four-lane arterials, collectors, and local roads. The research team obtained geometric, traffic, and crash data at treated locations in Washington and Texas. Unfortunately, a before–after study was not feasible because of the anticipated amount of data needed to obtain meaningful results. The limited availability of bicycle-volume data to estimate average daily bicycle traffic (ADBT) was a controlling factor in the feasibility of developing a before–after study with a sufficiently large sample size. Instead, the research team pursued a cross-sectional design, which used generalized linear models and generalized linear mixed models. Limitations regarding endogeneity and bias can be mitigated when developing the database by incorporating untreated sites (i.e., those without bicycle lanes) with similar characteristics to sites with the intervention (added bicycle lanes and reduced widths for lanes and shoulders) to the extent practicable so that a comparison between treated and comparison sites is as fair as possible. The research team applied propensity score methods to guide data collection and also for the statistical analyses to improve the balance between treated and comparison sites (Li et al. 2013, 2018; Banihashemi 2016). These methods are well documented for their applicability to causal effects estimation problems, as is the case in this report (Imai and Ratkovic 2015; Vermeulen and Vansteelandt 2015).

The results for Washington suggested reductions for fatal and injury, PDO, and bicycle crashes, but the large uncertainty associated with all these effects resulted in statistically insignificant CMF estimates at a 95-percent confidence level. The CMF for total crashes was either 1.31 or 0.79, depending on whether bicycle-lane additions were achieved by reducing shoulder widths.

The results for Texas were consistent with the directions and magnitudes of the analysis of Washington data, which was expected if both analyses estimate the same underlying CMF with the potential for biases well mitigated and differ only in sample size. In contrast to the Washington results, the larger dataset from Texas yielded CMF estimates with statistical significance at the 95-percent confidence level. With varying degrees of uncertainty, CMFs for total and fatal and injury crashes were found to range between 0.514 and 0.734. All CMFs for fatal and injury crashes and for total crashes were found statistically insignificant. The B/C ratio estimated for Texas two-lane undivided urban collectors and local streets was 16.61, assuming the ADBT does not increase after the installation of bicycle lanes.

CHAPTER 1. INTRODUCTION

Bicycle lanes are travel lanes dedicated to bicyclists along a street. Effective bicycle lanes promote a consistent separation between bicycles and passing motor vehicles. Implementing bicycle lanes may also benefit pedestrians because dedicated bicycle lanes shift some of the bicycles from adjacent sidewalks to the active travel way (American Association of State Highway and Transportation Officials (AASHTO) 1999; National Association of City Transportation Officials (NACTO) 2005). A bicycle lane can also help raise awareness among motorists that bicyclists are present and that they should be alert to these vulnerable users. The basic design consists of striping and signing. Additionally, some agencies apply colored pavement markings (e.g., green or red) to the bicycle lanes. In a constrained environment, however, an agency must consider reducing existing space previously dedicated to motor-vehicle use to accommodate bicycle lanes. This type of design is accomplished by reducing the width of motor vehicle lanes or shoulders as needed for the bicycle lane. This approach may result in a tradeoff between bicycle and motor-vehicle safety.

Reducing motor-vehicle lanes to provide for bicycle lanes is a common scenario within urban environments, but the shoulder may be either reduced or repurposed in suburban and rural environments. Figure 1 shows a bicycle lane in Tampa, FL, that replaces the right shoulder.

Figure 1. Photograph. Bicycle lane in Tampa, FL.

In a report to enhance active transportation, Litman (2017) suggested that space requirements per passenger-mile or kilometer, and therefore congestion impacts, vary depending on vehicle size, speed, occupancy, and number of interactions between bicycles and motor vehicles. Shy distance (the space between a vehicle and other objects) increases exponentially with speed. Litman observed that there is little impact in terms of congestion at busy roads with space for bicyclists, where bicycling occurs on a road shoulder, a wide curb lane (WCL), or a bicycle lane. An exception to this trend is at intersections, where bicyclist activity may delay vehicle-turning maneuvers (Litman 2017).

LITERATURE REVIEW ON BICYCLE LANES, LANE REDUCTION, SHOULDER WIDTHS, AND CRASH MODIFICATION FACTORS

This section presents a summary of the relevant literature on the safety effectiveness of bicycle lanes. There are three subsections: general evaluations of bicycle lanes, safety tradeoffs of cross-sectional element modifications, and benefit–cost studies.

General Evaluations of Bicycle Lanes

Based on journal reviews that considered data from Canada and northern Europe (Denmark, Finland, Germany, Netherlands, and Sweden), Thomas and DeRobertis (2013) concluded that construction of cycle tracks on busy streets in urban areas reduces the number of collisions and injuries. In this context, a cycle track refers to a facility or space exclusively dedicated to bicycles. Multilane roads are relatively rare in Europe compared to the United States; therefore, Thomas and DeRobertis further suggested that this approach be considered as a general guide for bicycle facilities on other facility types (Thomas and DeRobertis 2013).

AASHTO provides guidelines for restriping existing roads with bicycle lanes. These guidelines provide standards to be followed when considering tradeoffs between reducing the width of travel lanes and reducing the total number of motor-vehicle lanes. AASHTO (1999) recommends removal of onstreet parking and restriping WCLs. According to FHWA, safety improves when travel lanes are offset from curbs, lanes are better defined, and parking is removed or reduced. Adding bicycle lanes can often improve sight distance. Increasing the turning radii at intersections and driveways and restriping pavement extend pavement life because these measures laterally shift vehicle traffic, resulting in fewer vehicles driving on well-worn ruts (FHWA 2000).

The City of Burlington, VT, evaluated the addition of onroad bicycle facilities and identified a negative safety impact associated with the 1.0- to 1.2-m-wide bicycle lanes: an increase in the number of collisions from narrower motor-vehicle lanes (mainly sideswipe and head-on collisions resulting from motorists failing to maintain their lane) (NACTO 2005). A categorical analysis of before–after collision data indicated that an increase in bicycle-lane width coincided with a decrease in the proportion of rear-end collisions and single-motor-vehicle collisions and an increase in the proportion of turning collisions (NACTO 2005).

An evaluation by Hallett et al. (2006) looked at onstreet bicycle facilities added to existing roadways and investigated various factors affecting onstreet bicycle facilities using multiple regression. This study was based on nearly 3,500 observations of motorists passing bicyclists and over 4,000 observations of motorists unaffected by bicyclists at 24 sites across Texas. The study concentrated on adding bicycle lanes without increasing the overall roadway width (therefore, by reducing space for motor vehicles). The researchers identified 2,712 injury bicycle crashes between the beginning of 1998 and the end of 2001 in 13 counties in the Houston–Galveston area (12,500 mi^2), but the research did not result in meaningful findings, likely because of the lack of bicycle-exposure data combined with other site conditions. Researchers studied the lateral position of bicyclists at another set of locations where the bicycle lane was widened, but the analysis failed to identify many significant differences associated with candidate variables, such as traffic volume and motor-vehicle speed. The main product of Hallett et al.'s research was the development of the passing event model (PEM), which provides a metric of lateral position for bicycles and motor vehicles at passing events. Hallett et al. (2006) noted that PEM output is an indicator of operational suitability at a road segment. The PEM index was derived from the statistical analysis of more than 8,000 field observations in which Hallett et al. studied the lateral position of bicyclists and motor vehicles in relation to the roadway width and striping. Hallett et al. also reported that based on survey data, bicycle lanes provided higher operational and comfort levels under all conditions except for very low traffic volumes on exclusively residential streets. These findings were contrasted to wide outside lanes shared with motor vehicles. On wide outside lanes, bicyclists tend to ride extremely close to the curb, and motorists in the adjacent lane tend to encroach frequently, resulting in low operational and comfort levels (Hallett et al. 2006).

Park et al. (2015) evaluated the safety effects of adding a bicycle lane by using an observational before–after study and analyzing with empirical Bayes (EB) and cross-sectional methods. The study by Park et al. (2015) focused on urban arterial data from Florida. Park et al. used 10 yr of crash, roadway, and economic data (2003–2012). The study was based on a set of 227 road segments with a known installation date of bicycle lanes. A total of 2,437 crashes—1,358 before bicycle-lane installation and 1,079 after—were identified in this set of sites. Florida-specific full negative binomial safety performance functions (SPFs) were developed from another set of 517 roadway segments with similar characteristics. This analysis also incorporated socioeconomic parameters when using the EB procedure.

Park et al. (2015) developed the following EB-based crash modification factors (CMFs): 0.829 (0.029 standard error (SE)) for total crashes, 0.804 (0.039 SE) for fatal and injury (KABC) crashes, and 0.439 (0.083 SE) for total bicycle crashes. Additionally, Park et al. performed a sensitivity analysis on how EB bicycle-lane CMFs for total crashes varied among subsets of sites with different characteristics in their dataset. Park et al. further described these trends, although most confidence intervals for the partial-data CMFs overlapped (thus not providing sufficient statistical evidence of this variation). Park et al. identified the following trends:

- The addition of a bicycle lane corresponds to an increasing CMF as the traffic volume per lane increases.

- The addition of a bicycle lane corresponds to a decreasing CMF with increasing median widths.

When subdividing findings based on bicycle-lane width, the CMF was statistically significant at subsets as small as 38 sites.

Park et al. (2015) suggested that there could be a tradeoff when cross-sectional elements are reduced to add a bicycle lane. The apparent benefit (i.e., changes in total crashes) is larger when adding a bicycle lane.

Hunter et al. (1999) evaluated the safety benefits of a bicycle lane compared to a WCL. In this study, authors analyzed videotapes of nearly 4,600 bicyclists (2,700 riding in bicycle lanes and 1,900 using WCLs) in Santa Barbara, CA; Gainesville, FL; and Austin, TX. Bicyclists approached and rode through eight bicycle lanes and eight WCL intersections with varying speed and traffic conditions. Researchers found that WCL sites experienced more wrong-way riding and sidewalk riding than sites with bicycle lanes. Motor vehicles passing bicycles on the left encroaching into the adjacent traffic lane occurred significantly more often for locations with WCLs. More bicyclists obeyed stop signs at bicycle-lane sites, but the probability of having an unsafe situation was higher at bicycle-lane sites when the stop sign was disobeyed. Conflict severity did not change with the type of bicycle facility, and the majority of observed bicycle–motor vehicle conflicts were minor.

Lee et al. (2015) used generalized nonlinear models (GNMs) to develop CMFs that represented the effectiveness of changing lane width in Florida. Researchers observed crash-rate reductions when lane width was greater than or equal to 12 ft. However, results varied with the posted speed limit because of a strong correlation between lane width and speed limit. The relationship between lane width and operational speed is positive, as suggested by the *Highway Capacity Manual* (TRB 2016). A reduction in free-flow speed is expected with narrower lane width and lateral clearance. CMFs developed by Lee et al. indicated changes in the crash severity associated with changes in lane width. Lee et al. further recommended the use of GNMs instead of generalized linear models (GLMs) when the relation between crash rate and lane width is nonlinear.

Safety Tradeoffs of Cross-Sectional Modifications at Other Facilities

A Texas study by Dixon et al. (2016) examined the effects of adding a freeway lane while reducing the adjacent lane and shoulder widths. The authors determined that the construction of an additional freeway lane was associated with a 33-percent increase in fatal and injury crashes when a reduction in shoulder and adjacent lane width was necessary to accommodate the additional freeway lane (Dixon et al. 2016).

Chen and Tian (2012) conducted a statistical analysis of crash data from June 2007 to June 2011 on 304.5 mi of expressways and 56 interchanges in China. There were 2,935 traffic crashes at the study sites. Researchers considered only total crashes and fatal crashes in their evaluations. Chen and Tian observed that the crash rate at study locations was highest at a shoulder width of 12.3 ft, and this rate decreased with increasing shoulder width. The accident mortality rate was inversely proportional to shoulder width at interchanges. The research concluded that wider shoulders are associated with fewer crashes and lower mortality rates (Chen and Tian 2012). Bonneson and Pratt (2009) developed Texas-specific CMFs for rural segments. Their findings indicated major crash reductions for wider lanes, wider shoulders, and wider lane–shoulder combinations.

Gross et al. (2009) developed CMFs for several lane–shoulder combinations as part of a case-control study that evaluated low-cost safety strategies, including the reallocation of total paved width, based on data from Pennsylvania and Washington. Researchers obtained geometric, traffic, and crash data for the entire population of 52,000 mi of rural two-lane undivided road segments in Pennsylvania (1997–2001 and 2003–2006) and Washington (1993–1996 and 2002–2003). Gross et al. found an apparent tradeoff indicating a slight benefit to increasing lane width for a fixed total width. However, individual State results failed to indicate a clear tradeoff between lane and shoulder width for a fixed total width when compared with previous research.

Manuel et al. (2014) performed a cross-sectional study to analyze the safety implications of oversized collector roadways in Edmonton, Canada. Based on negative binomial SPFs, researchers found that lane width had a negative association with crashes, but the effect waned with increasing traffic along road segments. Manuel et al. found the safety of standard and oversized roads was worse than the safety of roads with uniform width. The conversion of oversized roads to standard width was associated with a crash reduction of 28.9 percent and an increase in traffic flow (Manuel et al. 2014).

Benefit–Cost Studies

Weigand et al. (2013) conducted a cost analysis of bicycle facilities in Portland, OR, and identified the following benefits of establishing bicycle lanes:

- Reduced frequency and severity of vehicle–bicycle crashes.
- Improved mobility for bicyclists.
- Reduction in speeds by motorized traffic.

Potential costs of bicycle facilities included the following:

- Restriping of streets to reduce lane/shoulder width and add bicycle lanes.
- Maintenance costs.
- Costs due to increased crashes resulting from narrower lanes/shoulder.
- Increased traffic delays.

Costs for bicycle facilities, however, were not easy to capture for several reasons. These costs often occur as part of larger multimodal roadway projects, and bicycle-related costs for these types of projects generally are not tracked separately (Weigand et al. 2013).

National Cooperative Highway Research Program Report 552 lists costs typically associated with onstreet bicycle facilities and discusses the common challenges with estimating the associated costs and benefits (Krizek et al. 2006). Costs may vary between States. Cost details were obtained from 40 States to create a database consisting of 1,747 records. States with the most cost information included Ohio (161), California (146), Minnesota (115), Massachusetts (104), and Wisconsin (101). States that had no information in the database were Delaware, Hawaii, Mississippi, Nevada, Pennsylvania, South Dakota, Tennessee, Utah, and West Virginia. The District of Columbia also had no information in the database.

The Pedestrian and Bicycle Information Center (2015) provides cost information for adopting certain techniques to provide narrow lanes. Reducing lane width by adding bicycle lanes costs at least $5,000 per mile. However, this cost varies widely based on the condition of the pavement and need for resurfacing improvements. Complete restriping of a street to reduce lanes or add bicycle lanes or onstreet parking costs approximately $5,000 to $20,000 per mile.

Improvements to the existing right-of-way keep the project cost to a minimum because this approach assumes no right-of-way acquisition cost or road-widening cost (Weigand et al. 2013). Onroad bicycle facilities are associated with lower costs if restriping the roadway to remove or narrow travel lanes is part of roadway repaving or reconstruction. Striping of travel lanes, shoulders, and bicycle lanes costs approximately $14,000 per mile for a 4-inch white solid line on one side of a lane. Restriping 1 mi to add bicycle lanes and reduce the number of traffic lanes costs approximately $20,000 to $48,000 per mile, depending on the number of old lane lines to be removed. Widening the road to add lanes as part of the roadway construction project has a reconstruction cost of approximately $287,000 to $300,000 per mile (Rivers and Associates, Inc. 2016). The benefit–cost (B/C) analysis should consider direct benefits to users from improved active transport conditions (i.e., improved road conditions for human-propelled transportation modes) and benefits to society from increased walking and cycling activity and reduced motor vehicle travel (resulting in more compact and multimodal community development) (Litman 2017).

SUMMARY

This chapter introduced the project and outlined the characteristics of the safety evaluation of introducing bicycle lanes while reducing other cross-sectional elements. A brief literature review described past research on the topic and similar evaluations regarding safety tradeoffs. This chapter presented cost elements found in the literature that will be used in the economic analysis in chapter 5. The next chapter outlines the study design and analytical methods implemented for the safety effectiveness evaluation in this project.

CHAPTER 2. STUDY DESIGN AND STATISTICAL METHODOLOGY

The research team evaluated multiple factors when designing this study. The study design had to account for the features of the available data and the ultimate purpose of the analysis.

Researchers identified potential data sources from which to gather key data elements for each improvement. Two basic designs for observational studies—cross-sectional and before-and-after—are frequently used in safety evaluations.

A strong study design can significantly boost the quality of the results; therefore, it is critical to closely examine all potential data sources, their characteristics, and available data elements. This examination should precede any data acquisition/collection and consider the needs for the analysis phase.

Safety studies are often limited to evaluations of observational data because randomization is not possible and true experiments, such as randomized control groups, are not feasible. Good observational studies rely on data from both treated and nontreated sites in a manner consistent with control-group experiments. A cross-sectional data analysis with no matching or control group is considered an inferior preexperimental design and is sometimes called a static-group comparison (Campbell and Stanley 1966). Likewise, if before–after data are analyzed with no control group, the quality of the design (one-group pretest–posttest design) is negatively affected. These types of preexperimental designs have a higher potential for biased results. This study, therefore, initially targeted a quasi-experimental design to the extent possible, such as a nonequivalent control group (or comparison group) design or a control series design (e.g., Campbell and Russo 1999; Campbell and Stanley 1966). However, when evaluating the tradeoff of a bicycle lane versus travel lane or shoulder widths, obtaining before–after data from multiple jurisdictions was deemed infeasible after reviewing potential data sources. Therefore, the research team developed a database for cross-sectional analysis. To incorporate comparison sites, team members adopted propensity score weighting (PSW) methods to minimize imbalances between covariates. More details about these types of adjustments are provided in this chapter.

STATISTICAL METHODS IN THE DATA-MANAGEMENT PROCESS

The data-management stage involves collecting and revising data from multiple sources, supplementing data where appropriate, concatenating variables across sources, and preparing data for statistical analyses. In response to actual data availability, team members refined datasets through data integration and data balancing.

Team members examined a variety of data sources identified in the feasibility study phase. One goal was to select candidate sites for study while balancing the features of treated and control sites. This database-development approach is consistent with the selected study-design methods.

Data Extraction and Integration

The research team used geographic information systems (GIS) tools to prepare, filter, and combine data containing multiple-source geolocation (typically in shapefile format). GIS tools allow the manipulation, combination, and display of data for different types of information, including crashes, road infrastructure, traffic volume, census tract, land use, and others.

Data Balancing

Data-matching and data-balancing methods are used to assist causal inference, which quantifies the impact of a treatment variable on a given response variable. Data matching is essentially a way to achieve data balancing in which each treated site is matched with at least one nontreated site. The result is a robust comparison of the response mean between groups of treated and nontreated sites while factors other than the one under evaluation are equally represented at all levels of the variable of interest. In this context, nontreated sites are also referred to as control sites. The main principle behind this effort is to identify untreated locations similar in their covariates to the treated locations so that the contrast by the response variable is implicitly controlled for other covariates that could have an impact on the response variable. In cases where treated data are scarce, it is common to use all identified treated sites and then match control sites using sampling techniques on a wider sampling frame of candidate control sites. Matching treated and control sites is based on the covariates identified to covary with the treatment variable.

An important step in data matching is validating the matched data, which is also called the quality-control step. Matched data are validated by examining the mean of the covariates across the treatment and control groups. A balanced dataset is one in which the means are almost identical, implying that any observed difference between the treatment and control groups in the response variable is from the effect of treatment. Crashes on selected untreated locations are then used as a proxy to estimate the counterfactual crashes on the treated locations—that is, the crash frequency that would have been observed if the treatment had not been applied. The quality of the matched dataset can be increased by matching one treated site with several control sites. The epidemiology literature recommends selecting between two and four control units per one treated unit (Lewallen and Courtright 1998; Linden and Samuels 2013).

Propensity Score Methods

More analytical approaches to guide the data-matching phase are based on propensity scores (PSs). The PS is a metric of similarity between covariates from the cases and can be estimated using parametric or nonparametric tools, such as logistic regression or random forest analysis (Guo and Fraser 2014; Jovanis and Gross 2007; Sasidharan and Donnell 2013, 2014). Under this framework, the PSs of treatment cases and their corresponding control cases are estimated and compared.

In the case of binary logistic regression as a basis for PS estimation, figure 2 shows the definition of the conditional probability of a site receiving treatment *T*.

$$P(T_i|X_i) = \frac{e^{\alpha_i X_i}}{1 + e^{\alpha_i X_i}}$$

Figure 2. Equation. PS definition as a logistic function of covariates.

Where:

$P(T_i|X_i)$ = PS denoting the probability of the site i receiving the treatment T.
T_i = treatment status of the site i, which takes binary values {0, 1}.
X_i = vector of covariates that covary with the treatment presence.
α_i = vector of coefficients through the binary logistic regression.

In a balanced sample, the distribution of PSs is expected to be similar for treated sites ($P(T_A|X_A)$) and control sites ($P(T_B|X_B)$). An examination of these differences at various stages of data collection can be used to direct data collection at additional control sites to improve the balance of the dataset.

An alternative to propensity score matching (PSM) is PSW. In PSW, the PS is still used to balance two or more partitions of the data by the variable of interest (i.e., treatment or control). In contrast with PSM, balance is achieved by defining appropriate weights for each unit of analysis so that they represent an underlying target population of sites. Data are weighted based on the probabilities of being in either the control or the treatment group, and the selection of the weights defines the target population (Olmos and Govindasamy 2015). If all weights are equal, then the database is implied to be a simple random sample from the larger pool of sites from which data were collected. That pool of sites is then implicitly defined as the target population. However, more flexible definitions of the target population are possible by using appropriate weights; several of these flexible definitions can be found in the statistical literature (Olmos and Govindasamy 2015). In some cases, it makes sense to define a theoretical population most suitable for inference. The definition of the weights also determines quantities that can be estimated, including the average treatment effect and the average treatment effect among treated cases, control cases, and the evenly matchable cases.

STUDY DESIGN

Based on the findings of the feasibility study, team members collected and assembled data for a cross-sectional estimation of the CMFs for adding a bicycle lane with a reduced lane/shoulder. Control sites were also added to strengthen the design. Data collection required a metric that represented bicycle exposure so that bicycle traffic could be included in the estimation of exposure. Ideally, this metric would be a direct bicyclist volume count, which was the case in some but not all sites. In cases where such a direct metric was not available, direct demand estimates of bicyclist volumes from appropriate models were incorporated instead. The research team also implemented strategies to balance covariates accordingly.

Initially, the research team used matching methods (e.g., those found in Stuart (2010)) based on the PS to obtain control units to ensure similar covariate distributions to the treated sites. However, as documented later in this report, this strategy is most suitable with large pools of potential sites (e.g., databases from Texas) where multiple potential controls are available for matching. This approach is challenging for small pools, such as the Washington database after filtering out unfeasible locations. The research team therefore adopted the PSW framework. The target population was set to be the overlap between the treated and control populations, as proposed by Li et al. (2018). Under this scheme, the target population is the set of all sites that have comparable chances to be in either the treatment or the control group. This approach effectively curbs the undue influence of the following two subsets of sites when estimating the average treatment effect of the countermeasure:

- Control sites whose characteristics make them unlikely to be candidates for the treatment.
- Treated sites with unusual characteristics for which no comparable control sites are represented in the data.

An additional advantage of this choice of target population is a desirable small-sample exact-balance property, as demonstrated by Li et al. (2018). Corresponding weights minimize the asymptotic variance of the weighted average treatment effect within their class of weights (Li et al. 2018).

The research team used PSs to direct balancing of datasets during data collection, and PSW was implemented in the subsequent analysis, which used the PSs obtained from final datasets.

DATA-ANALYSIS METHODS

Empirical analyses were conducted using the statistical methods appropriate to the characteristics of the assembled datasets. The research team used appropriate GLM variants (e.g., negative binomial, Poisson-lognormal mixture, logistic-lognormal mixture).

Generalized Linear Regression Analysis With PSM or PSW

The predictive methods described in AASHTO's *Highway Safety Manual* (2010) are based on cross-sectional statistical models named SPFs. These models estimate the long-term expected crash frequency from multiple sites with similar characteristics. In this framework, it may be tempting to assume that the effect of a countermeasure can be estimated simply by comparing the counterfactual crash frequencies between treated and nontreated sites.

That approach is not completely without value, but a latent risk of this comparison is that the comparison group—the basis for the SPF—may not necessarily be representative of sites having the treatment under study. One way to reduce that risk is to develop the SPF based on a probability sample of the types of sites of interest. Another possibility is to develop the SPF from a complete population of sites (e.g., all sites in a State database inventory) whenever feasible. Such alternatives, however, are not possible or practical in every case. By using PS-based methods (PSM or PSW), the effect of a treatment can be studied by employing sites with treatment and matching untreated sites in a resulting dataset that has characteristics mimicking those expected from a randomized sample (Rosenbaum and Rubin 1984). PS methods ensure that covariates in the treated and untreated subsets of data are roughly independent, which should result in a nearly unbiased, nonconfounded estimate of the effect of interest. The effects of selection bias otherwise present in developing the cross-sectional dataset are thus mitigated.

Mixed-Effects Models

Within the frame of GLM methods, a distinction can be made between models with fixed effects, random effects, and mixed effects.

Commonly, the coefficients obtained from GLMs can be thought of as fixed effects. The variables corresponding to fixed effects are implied to have time-invariant effects (e.g., roadway design elements). The model coefficients are estimated and interpreted as metrics of underlying parameters from a latent data-generating process.

Random-effects models estimate the effects of factors that are deemed the observed realizations of a random variable. It is typically not of interest to quantify how the response variable shifts with the observed realizations in the dataset but rather to account for the impact of such variations in the model. The simplest analogy of random effects in a GLM is the use of blocking in analysis of variance designs. The effect of each block is not the focus of the analysis; however, the effect is of interest to account for the variability explained by the blocking to quantify the variability explained by the independent variable of interest.

Mixed-effects models are models that include both fixed and random effects (Pinheiro and Bates 2000). Generalized linear mixed models (GLMMs) approach the analysis of repeated measures cross-sectional data by including a random effect per every unit of data aggregation (i.e., the blocking units in the data, such as individual study locations with more than one datum in the analysis). Independent of random effects, the model estimates fixed effects for the treatment and any additional fixed-effects covariates. As in GLM methods, an appropriate link function can be specified to permit the modeling of count-data distributions, such as Poisson and negative binomial, that are applicable to crash data.

As described in the previous section, using PSM in the data-collection stage can produce a more robust dataset, and the resulting PS can be incorporated through PSW in the analysis stage, including the use of mixed-effects models.

CMF Estimation

In most cases, using regression models to estimate the influence of a dependent variable consists of extracting a single parameter estimate and its SE from the analysis after accounting for additional variability in the data through model covariates and an appropriately modeled error distribution. Single-parameter estimation, however, is a challenge for the treatment of interest in this report—estimating a shift in safety effectiveness when adding a bicycle lane and reducing other cross-sectional elements. The treatment of interest involves changes in at least two cross-sectional elements; therefore, the CMF estimation in this case involves two or more model parameter estimates if the model accounts for individual cross-sectional elements explicitly. To estimate the uncertainty of this compound CMF, the research team implemented the methods outlined in the next sections. These methods leverage the asymptotically multivariate normal distribution expected from multiple variable model estimates obtained from maximum likelihood estimation (Booth and Hobert 1998; Morrell et al. 1997; Wackerly et al. 2008).

CMF Estimates for Interventions with Multiple Effects

Best-fit models are generally not expected to produce a single coefficient estimate of the change in safety performance in a facility when a bicycle lane is accommodated while modifying lane and shoulder width. Initial and final widths for bicycle lanes, motor-vehicle lanes, and shoulders would influence the joint estimate of such modification. Additionally, literature suggests that the construction of a bicycle lane tends to attract more bicycle traffic (Litman 2017; Manuel et al. 2014).

There are analytical options to estimate the impact of multiple roadway features changing (i.e., bicycle lanes, motor-vehicle lanes, shoulders, and bicycle volumes). The following sections describe the estimation procedures the research team implemented.

Scenario-Based CMF Estimation

The coefficient estimates from each of the best models involving the variables of interest can be used to develop crash predictions via appropriate linear combinations for select scenarios. Linear combinations produce predicted crashes for both the before condition (e.g., no bicycle lane and baseline average daily bicycle traffic (ADBT)) and the after condition (e.g., bicycle lane, reduced lanes and shoulders, and increased ADBT). A contrast between the predictions then produces the estimated CMF. Producing the CMF estimate for the scenario is straightforward; however, producing the corresponding SE is a more complicated but feasible task. For a scenario with variable vectors X_B and X_A representing the before and after conditions of the safety influential variables in the model, respectively, and the inverse maximum likelihood information matrix **Σ**, figure 3 gives the SE for the contrast (i.e., CMF estimate) (Johnson and Wichern 2007; Wackerly et al. 2008).

$$SE(\log CMF) = \sqrt{(X_A' - X_B') \times \Sigma \times (X_A - X_B)}$$

Figure 3. Equation. SE for contrast in log scale.

The potential issue with applying figure 3, however, is that it requires the analyst to exercise some judgment when determining the levels for all covariates in each scenario. There is an inherent risk of bias when determining values for the covariates, even when selecting values that are within the range covered in the data used to fit the model. It is possible that the specific combination of variable or covariate values may not be present in the dataset; therefore, such estimates carry an increased (but undetermined) uncertainty.

CMF Estimation From Contrast Between Crash Expectation and Counterfactual

To curb the risk of bias associated with defining covariate levels for contrasts, the research team opted for a third approach: construct a counterfactual for each observation in the dataset and then contrast the crash prediction to the corresponding counterfactual to estimate the safety effectiveness for each unit of analysis. The CMF estimate is the overall average of all the estimates of safety effectiveness. Using this approach, analysts need not assume the levels of covariates to construct a scenario for contrasting estimates; only key variables need to be modified to construct the counterfactual.

SE estimates are derived from the model uncertainty estimation reflected in the inverse-information matrix. Figure 4 gives the SE of the average effect of the intervention (i.e., the SE of the CMF estimate) for design matrix *A*, counterfactual matrix *B*, inverse-information matrix Σ, and vector of weights *w* (derived from PSs, defining the overlap population).

$$SE(\log CMF) = \sqrt{\boldsymbol{w}' \times \{(\boldsymbol{A} - \boldsymbol{B}) \times \boldsymbol{\Sigma} \times (\boldsymbol{A}' - \boldsymbol{B}')\} \times \boldsymbol{w}}$$

Figure 4. Equation. SE for the average effect of the intervention.

In figure 4, as well as in the calculation of the CMF estimate, the weights are defined as the overlap weights from the PS analysis. This definition ensures that the estimate represents sites with comparable chances to be in any of the treatment-level groups (either with or without a bicycle lane). Additionally, the formulation in figure 4 explicitly considers the correlation among crash expectations and counterfactual predictions. Such estimates are not statistically independent because they are all derived from common model estimates that are correlated. The estimate of the SE in figure 4, therefore, is reflective of the uncertainty of the model estimates equivalent to the uncertainty expressed in the SE of a single coefficient from a regression model.

SUMMARY

This chapter described the statistical methodology, analyses methods, and tools that the research team used to perform the work in this project. This chapter presented and discussed the challenges associated with the evaluation and the critical step to develop a database with bicycle volume estimates. The rationale for a cross-sectional study design was presented, and reasons the PS methods were appropriate to reduce the risk of biased estimates in cross-sectional designs were discussed. This chapter also outlined how PS methods can be used to guide database development so that the resulting databases are naturally balanced in their key covariates. Finally, this chapter outlined statistical analysis methods to develop statistical models of crashes to be used for developing CMFs of interest. The chapter ended with a discussion of why the safety effectiveness estimate may not be captured by a single coefficient in the crash models and how additional techniques based on mathematical statistics can be applied to develop CMF estimates. Chapter 3 examines the data-collection effort for Washington and Texas in more detail.

CHAPTER 3. DATA COLLECTION AND INTEGRATION

Researchers reviewed potential datasets from multiple States (Florida, Oregon, Texas, and Washington). The limited availability of locations with actual bicycle counts or a potential for estimation (e.g., through the crowdsourced database Strava) influenced the decision to narrow the evaluation to the two States with the most promise to develop the dataset for analysis: Washington and Texas. Researchers chose these two States because of the number of potential locations and the quality of the bicycle exposure estimates that could result. To develop the database, researchers collected the following data elements:

- Bicycle and traffic counts.
- Bikeway facility type.
- Multiple roadway design elements (e.g., functional class, number of lanes, and lane and shoulder widths).
- Posted speed limit.
- Crash data (e.g., location, year, type, and severity).

This chapter discusses the data-collection and data-integration steps.

WASHINGTON

Bicycle-Count Data

Researchers obtained bicycle-count data from the Washington State Department of Transportation Pedestrian and Bicycle Data Collection portal (Van Schalkwyk and Washington 2008). The Washington State Documentation Project collects bicycle- and pedestrian-usage data in cities throughout the State. The documentation project uses a data-collection protocol similar to and consistent with the National Bicycle and Pedestrian Documentation Project (NBPDP). Counters and volunteers collected data from permanent and temporary sites. Data were gathered from a network of city staff, bicycle club members, and other volunteers to collect volume counts and document the data using the standardized NBPDP process. Data were collected from the following cities:

- Anacortes.
- Bainbridge Island.
- Battle Ground.
- Bayview.
- Bellevue.
- Bellingham.
- Bothell.
- Bremerton.
- Burien.
- Burlington.
- Concrete.
- Ellensburg.
- Everett.
- Federal Way.
- Ferndale.
- Gig Harbor.
- Issaquah.
- Kelso.
- Kenmore.
- Kent.
- Kirkland.
- La Conner.
- Lake Forest Park.
- Lakewood.
- Longview.
- Lyman.
- Lynden.

- Mercer Island.
- Milton.
- Mount Vernon.
- Mountlake Terrace.
- Oak Harbor.
- Olympia.
- Orting.
- Parkland.
- Pasco.
- Pullman.
- Puyallup.
- Redmond.
- Renton.
- Richland.
- Seattle.
- Sedro-Woolley.
- Shoreline.
- Skagit County (unincorporated).
- Spokane.
- Spokane Valley.
- Sumner.
- Swinomish Indian Tribal Community Reservation.
- Tacoma.
- Tukwila.
- University Place.
- Vancouver.
- Vashon Island.
- Walla Walla.
- Wenatchee.
- Yakima.

In this project, the research team examined a dataset of historical counts from 398 locations that included segments, intersections, and shared-use paths (trails). Washington State department of transportation (DOT) estimates ADBT data for the following times:

- 7–9 a.m.
- 4–6 p.m.

To estimate the overall ADBT, researchers used the total counts throughout the day.

Bikeway Facility Type and Roadway Data

The research team collected roadway-characteristic data from locations with bicycle counts. The team used data from the second Strategic Highway Research Program's Roadway Inventory Database (RID) (Iowa State University 2014) and Google Earth (2019). Researchers then combined the count stations with RID links. Only 55 of 398 count locations overlapped with a link, including all geometry and traffic characteristics. In addition to these segments, researchers verified that the locations extended at least for the length of the corresponding GIS links in the database. Research team members included additional sites that corresponded to the extension of the adjacent segments (represented by the GIS digitized vector links) after verifying that the cross section extended uniformly into those additional segments. This additional data assessment increased the number of segments in the Washington database to 87.

Crash Data

Researchers integrated the resulting links with crash data and identified all crashes within 100 ft of the roadway segments included in the database. Data included crashes located at sites that were both on system and off system. Off-system crashes were only available from 2010 to 2013 because of a change in the data-collection method. The research team therefore used 4 yr of crash data. After removing intersection- and driveway-related crashes, 2,262 crashes remained for analysis. Table 1 shows the descriptive statistics of the resulting database.

Table 1. Descriptive statistics of segments in Washington.

Variable Name	Minimum	Maximum	Mean	Standard Deviation
Segment length (ft)	65	3,766	572.01	562.40
Number of through lanes	1	6	2.72	1.14
Left-turn lane	0	2	0.29	0.50
Right-turn lane	0	1	0.21	0.41
Lane width (ft)	9.16	14.86	11.19	2.20
Shoulder width (ft)	0	11	1.40	2.89
Bicycle-lane width (when present) (ft)	4	6	1.65	2.60
Year	2010	2013	2011.50	1.12
ADBT (7–9 a.m.)	0	279	28.28	55.59
ADBT (4–6 p.m.)	0	468	34.61	68.81
ADBT	4	739	63.20	119.31
AADT	1,000	160,504	22,894.99	28,240.16
VMT	0.54	152.57	11.13	19.41
Total crashes	0	83	6.50	11.35
Fatal crashes	0	5	0.02	0.29
Incapacitating injury crashes	0	10	0.10	0.91
Nonincapacitating injury crashes	0	17	0.47	1.58
Possible injury crashes	0	26	1.86	3.86
Fatal and injury crashes	0	41	2.45	5.13
Bicycle crashes	0	9	0.07	0.56

AADT = annual average daily traffic; VMT = vehicle miles traveled.

Four other variables are not shown in table 1 that were binary indicator variables for the presence of bicycle lane, bus lane, parking lane, and two-way left-turn lane.

TEXAS

Bicycle-Count Data

In this project, the research team used estimated bicycle-count data from the crowdsourced database Strava (Strava 2018). The project applied the direct-demand models developed in Texas Department of Transportation Project 0-6927 to estimate the bicycle counts from the crowdsourced database (Turner et al. 2019) (table 2).

Table 2. Direct-demand models (Turner et al. 2019).

Strava Definition of Roadway Functional Class	Direct-Demand Model
Highway, primary	$ADBT_i = 63 \times e^{(0.038 \times ADBT\ Strava\_i + 0.002 \times \text{Household Density\_}i)}$
Highway, secondary	$ADBT_i = 13 \times e^{(0.038 \times ADBT\ Strava\_i + 0.002 \times \text{Household Density\_}i)}$
Highway, tertiary	$ADBT_i = 22 \times e^{(0.038 \times ADBT\ Strava\_i + 0.002 \times \text{Household Density\_}i)}$
Highway, residential	$ADBT_i = 17 \times e^{(0.038 \times ADBT\ Strava\_i + 0.002 \times \text{Household Density\_}i)}$
Highway, path	$ADBT_i = 72 \times e^{(0.038 \times ADBT\ Strava\_i + 0.002 \times \text{Household Density\_}i)}$
Cycleway	$ADBT_i = 62 \times e^{(0.038 \times ADBT\ Strava\_i + 0.002 \times \text{Household Density\_}i)}$
Footway	$ADBT_i = 28 \times e^{(0.038 \times ADBT\ Strava\_i + 0.002 \times \text{Household Density\_}i)}$

The 0-6927 Texas project collected count data from 124 roadway segments in 11 cities. Data can be visualized and queried from the Texas Bicycle and Pedestrian Count Locations Database (TTI 2019). These data were collected during different times from 2016 to 2017.

Turner et al. (2019) integrated the site counts with the Strava sample and developed direct-demand models to estimate the average annual daily bicycle counts. The ADBT model is constructed using the three most important variables determined to be significantly associated with bicycle use:

- Strava sample counts.
- Type of roadway functional class, based on OpenStreetMap definitions (OpenStreetMap® 2004).
- Density of high-income households in the given census block group, collected via the American Community Survey (U.S. Census Bureau 2018).

The mean absolute prediction error or error margin for the ADBT model was 29 percent. The overdispersion parameter was 1.147.

Bikeway Facility Type and Roadway Data

The research team used data on existing bikeway facilities obtained from TxDOT (TxDOT 2020). Researchers used the acquired shapefiles and selected segments with bicycle lanes and without bicycle lanes.

Researchers combined the selected segments (with and without an onstreet bicycle lane) with Texas roadway inventory and crowdsourced data to estimate bicycle counts on these segments. As a result, 5,473 segments were identified and included in the initial database. Table 3 shows the descriptive statistics of these segments.

Table 3. Descriptive statistics of segments in Texas.

Variable	Minimum	Maximum	Mean	Standard Deviation
Number of lanes	1	8	3.03	1.27
Surface width (ft)	14	96	35.17	15.49
Roadbed width (ft)	18	131	39.17	20.96
Inside (left) shoulder width (ft)	0	18	0.99	2.81
Outside (right) shoulder width (ft)	0	20	1.67	4.74
Median width (ft)	0	312	3.26	14.34
Segment length (ft)	5.28	34,473.12	4,509.64	3,903.63
Land area (m^2)	0	106.1693	2.26	5.94
Number of households with income >$200,000	0	1,611	83.37	141.64
AADT 2015	54	107,121	12,435.06	14,390.35
AADT 2016	10	92,462	13,001.15	14,756.86
Total annual number of Strava users	0	32,085	671.55	1,967.98
Annual average daily Strava users	0	87.90411	1.84	5.39
ADBT (estimated)	0	1,135	38.67	28.26
Total crashes, 2015	0	128	2.05	6.38
Total crashes, 2016	0	79	1.12	3.56
Total crashes, 2017	0	76	1.09	3.44
Total crashes, 2018	0	31	0.50	1.61
All crashes	0	266	4.77	14.40
Bicycle crashes, 2015	0	4	0.01	0.14
Bicycle crashes, 2016	0	3	0.01	0.11

Variable	Minimum	Maximum	Mean	Standard Deviation
Bicycle crashes, 2017	0	2	0.01	0.11
Bicycle crashes, 2018	0	1	0.00	0.05
Bicycle crash, all	0	6	0.04	0.25
Fatal crashes, 2015	0	3	0.01	0.11
Fatal crashes, 2016	0	2	0.01	0.09
Fatal crashes, 2017	0	1	0.00	0.07
Fatal crashes, 2018	0	2	0.00	0.06
Suspected serious injury crashes, 2015	0	4	0.05	0.27
Suspected serious injury crashes, 2016	0	4	0.03	0.19
Suspected serious injury crashes, 2017	0	4	0.03	0.21
Suspected serious injury crashes, 2018	0	3	0.01	0.12
Fatal and suspected serious injury, all	0	10	0.16	0.60
Nonincapacitating injury crashes, 2015	0	18	0.28	0.96
Nonincapacitating injury crashes, 2016	0	11	0.15	0.58
Nonincapacitating injury crashes, 2017	0	9	0.13	0.53
Nonincapacitating injury crashes, 2018	0	4	0.06	0.28
Possible injury crashes, 2015	0	33	0.43	1.57
Possible injury crashes, 2016	0	27	0.23	0.91
Possible injury crashes, 2017	0	25	0.23	0.91
Possible injury crashes, 2018	0	14	0.10	0.48
Property-damage-only crashes, 2015	0	115	1.75	5.65
Property-damage-only crashes, 2016	0	64	0.96	3.13
Property-damage-only crashes, 2017	0	60	0.92	3.00
Property-damage-only crashes, 2018	0	24	0.43	1.42

AADT = annual average daily traffic.

Additionally, four factor variables were included in the database: highway system (ranging from interstates to local roads), functional classification (similar to highway system but it makes a rural versus urban distinction), surface treatment type (including various levels of asphalt and hydraulic concrete), and shoulder type (including paved, gravel, and unpaved shoulders).

Crash Data

After identifying the segments of interest, the research team identified crashes that occurred on selected roadway segments. Bicycle counts were estimated for a period between July 2016 and June 2017; therefore, the research team selected crash data from 2015–2018 under the assumption that the bicycle lane remained present at the selected locations 1 yr before the data collection. Researchers used a geolocation buffer to identify the segment crashes. Team members then applied filters to remove crashes from adjacent locations. After identifying crashes corresponding to the facilities under study, all intersection-related crashes were removed from the database before analysis.

The assembled database described in table 3 was further filtered before subsequent analysis. For example, although some locations with eight lanes (i.e., four in each direction) were identified, these were removed from further analysis because there were very few of them. Additionally, some segment lengths were recorded around 7.0 mi, which is highly improbable in an urban environment. Locations with other uncommon conditions were represented in reduced numbers, which made it impractical to include them in this study. From an initial set of 5,473 segments with ADBT estimates available, researchers identified and discarded segments that had no bicycle lanes but had very wide two-lane cross sections or more than five lanes. Non-bicycle-lane segments with curbed medians were also discarded because the scarcity of this condition among the sites with bicycle lanes makes the comparison unrealistic. A total of 1,520 segments were discarded in this step. These segments were mostly city streets or local roads. Inspection of these discarded segments showed that the wider lanes were likely used as parking space in most instances. The research team decided to focus the evaluation on cross sections with two or four lanes because non-bicycle-lane segments with these conditions were most comparable to those that included a bicycle lane.

Further inspection of the reduced subset of segments indicated that facilities labeled as including bicycle lanes were facilities with wider travel lanes in general, sometimes including space for parking but without bicycle lanes. The research team decided to manually inspect a random sample of these locations to investigate a reliable way to identify and discard mislabeled non-bicycle-lane segments. A sample of 600 segments was collected and analyzed to develop a probability model for the confirmed presence of bicycle lanes. A cross validation of the model indicated an accurate classification rate of 89 percent. The model was then used to produce estimated probabilities of bicycle lanes over the entire dataset. The model and the reported facility type agreed for 70 percent of the dataset. For the remaining 30 percent, the locations indicated as having bicycle lanes but unlikely to have them, according to the model, were relabeled accordingly. In a last quality-control step, the research team verified that this procedure correctly identified locations previously mislabeled for a select subset of sites in the refined database, now containing 3,622 segments for analysis.

SUMMARY

This chapter documented the process for selecting the States for evaluation, choosing the data elements, and collecting the data for the safety evaluation of bicycle lanes. Summary statistics were presented for the two databases developed: one for Texas and one for Washington sites. The chapter presented a discussion of the reasons why data from both States could not be merged into a single database, namely because of significantly different sample sizes and data structure.

CHAPTER 4. SAFETY EFFECTIVENESS EVALUATION

This chapter describes the statistical evaluations of the datasets described in the prior chapter and the obtained CMF estimates for adding bicycle lanes as a safety improvement.

MODELING PROCESS

The research team used entropy metrics (Akaike information criterion and Bayesian information criterion) to guide model development. In each case, the research team found the best fitting model for the response variable of interest, namely crash frequency by severity and type, as shown in table 4.

Table 4. Response variable and model specifications.

Response Variable	Model Specifications Considered
Total crash frequency	GLM and GLMM
KABC crash frequency	GLM and GLMM
Bicycle crash frequency	GLM and GLMM

Researchers fit models as outlined in table 4 after sorting the data in a sensible way (i.e., by State, functional class, number of lanes). However, sample sizes in Washington were greatly reduced for these subgroups. This reduced subsample generally increased the SE of estimates. The research team then compared the coefficient estimates of key variables from the subgroup models (regardless of their statistical significance) and determined that in general the trends agreed (i.e., same order of magnitude and same direction of the effect when statistically significant).

No concerns were raised by the comparison of the subgroup models; therefore, the research team decided to fit an overarching model with all the subgroups, carefully testing for interactions among key variables that could indicate differences in safety performance by the subgroups.

ANALYSIS BY STATE

The research team performed the analysis separately by State, considering different variables, different number of locations, different geographic locations or sites, and differences in origin of bicycle-volume estimates. Differences in bicycle volumes were most influential in this analysis. The dataset from Washington offered actual counts for multiple years at each location under study, whereas bicycle volumes were estimated for the Texas dataset. By separately performing the analysis, the research team could identify potentially diverging trends and levels of accuracy between the two analyses. The team would then be able to assess how the two-volume estimates impact the evaluation of safety effectiveness at these facilities. The following sections document the analyses and results by State.

Washington

The research team prepared the dataset for two exploratory analyses: facilities with two lanes and facilities with four lanes. For each subset of sites, the team developed a PS model for a bicycle lane. Researchers then developed overlap weights from the PS values, as proposed by Li et al. (2018). These weights were defined to represent the population of sites in the overlap of the two subsets. Plots in figure 5 demonstrate the balancing effect of this procedure on average lane width, a key variable in the evaluation.

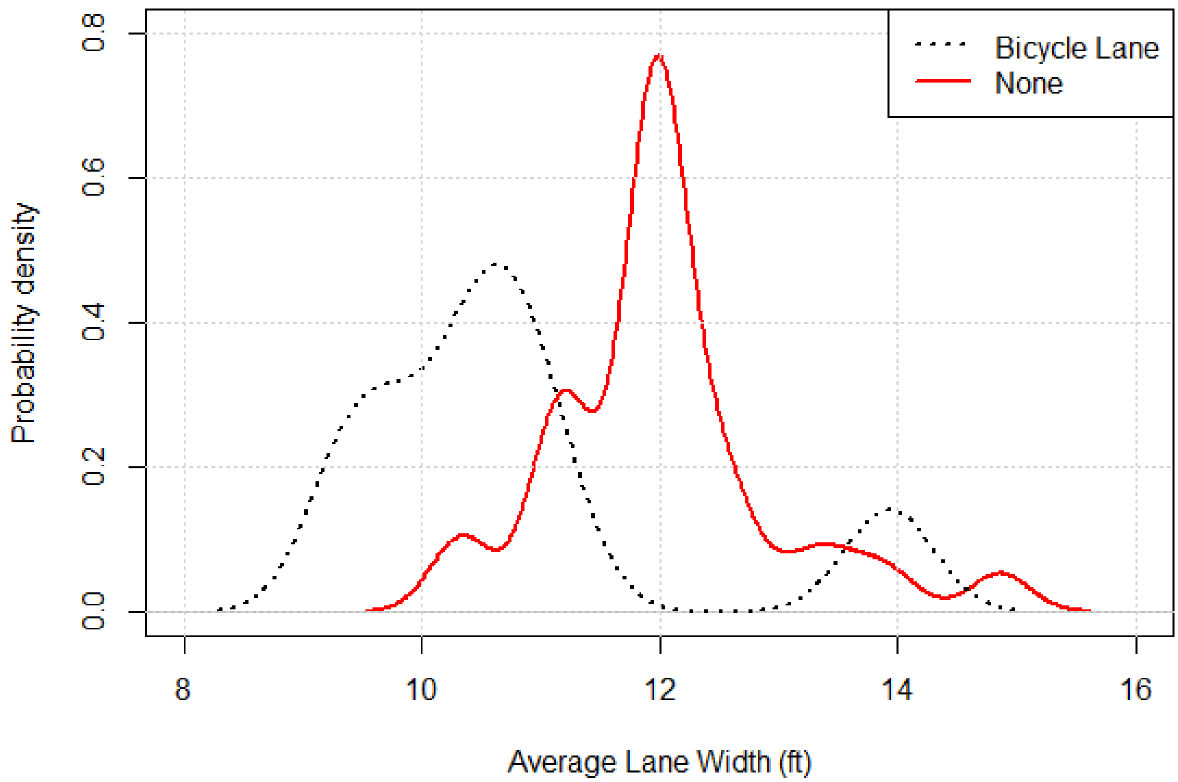

Figure 5. Graph. Average lane width unweighted distributions by presence of bicycle lane in Washington (two-lane segments).

Figure 5 shows an imbalance in lane width between the sites with and without bicycle lanes in Washington. Sites with bicycle lanes have narrower lanes than sites without bicycle lanes. In contrast, figure 6 shows how the application of PS weights balances the distributions, resulting in a more comparable contrast between the two data subgroups. Figure 6 also shows the overlap distribution line. The overlap distribution is the distribution of sites with characteristics that make the sites nearly equally likely to be in either the treated group or the control group. The overlap distribution, therefore, is the population of inference when applying PS weights.

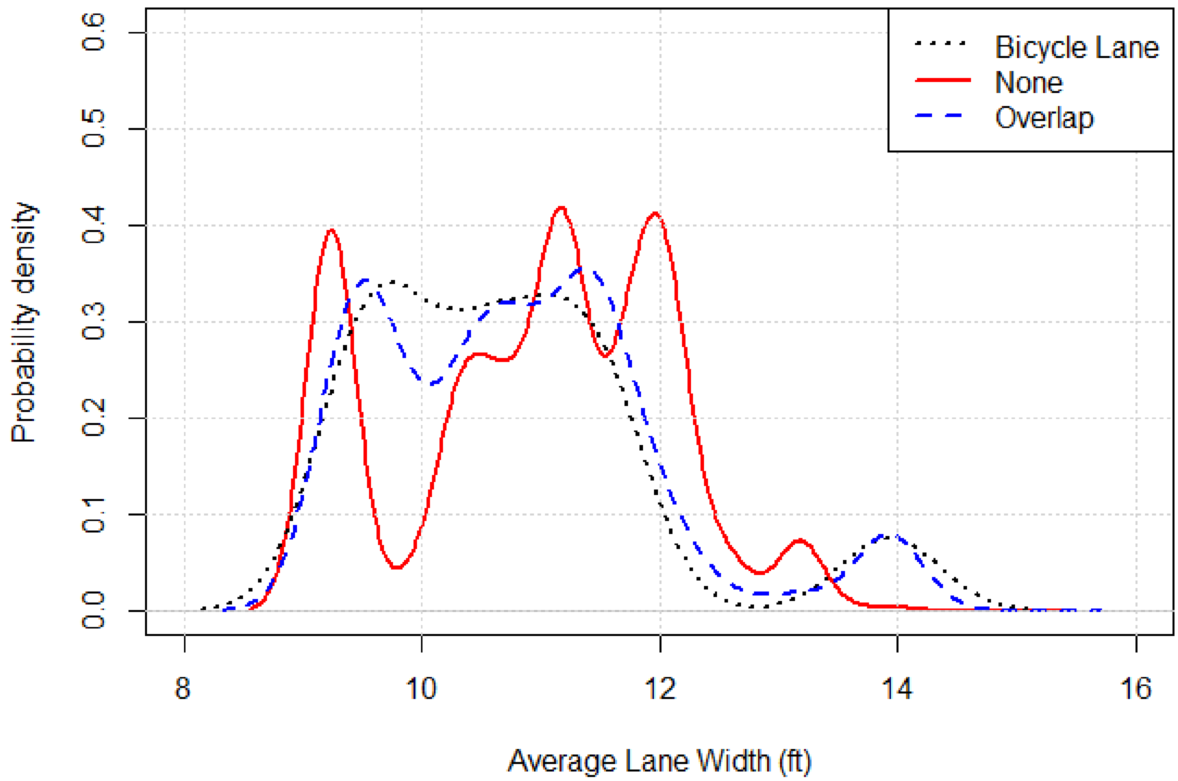

Figure 6. Graph. Average lane width weighted distributions by presence of bicycle lane in Washington (two-lane segments).

The reduced dataset for Washington precluded the research team from performing analyses in relatively uniform subsets of data. Instead, the research team performed an analysis on the complete dataset, as summarized in table 1.

Researchers performed the analysis without aggregating all years because data have yearly granularity in bicycle volumes. This approach allows leveraging year-by-year variability in the data and provides an increased number of degrees of freedom for estimating the SEs for the maximum likelihood estimates. Each of the 87 segments with bicycle volume in Washington had approximately 4 yr of crash and volume data. There were 343 segment-periods available for analysis.

Data-Analysis Results

Table 5 presents the results from the best-fitting models. As described in chapter 3, statistical models were fit for four crash responses defined in this dataset. The model for this analysis was a GLMM with a Poisson-lognormal mixture. This approach models site-to-site variability as a lognormal distribution, whereas the crashes within a site are modeled as a Poisson variable. Any Poisson overdispersion present in the data, therefore, is captured in the variability of the random effects. An estimate of Poisson overdispersion can then be constructed and reported, analogous to the dispersion parameter in the negative binomial distribution.

Table 5. Coefficient estimates for crash prediction models in Washington (*n* = 343).

Parameter	All Crashes		Fatal and Injury Crashes		Property-Damage-Only Crashes	
	Coefficient	Standard Error	Coefficient	Standard Error	Coefficient	Standard Error
Intercept (β_0)	−2.073**	0.6367	−3.336	2.0939	−4.7504**	1.631
Length in miles (α_1)	0.6843***	0.1761	0.5983*	0.2363	0.6864***	0.1768
AADT (α_2)	0.8518***	0.2129	0.4168~	0.2243	0.7046***	0.1754
Scaled (ADBT/AADT) (β_1)	—	—	−0.4977	0.3238	−0.0059	0.1428
Bicycle lane (β_2) (0 = no, 1 = yes)	0.3534	0.3202	0.2591	0.3841	−0.122	0.2865
Number of lanes (β_3)	0.1999	0.1304	—	—	—	—
Shoulder width (β_4)	−0.0644	0.1148	—	—	—	—
Presence of right-turning lane (β_5)	0.8779**	0.3226	1.3081**	0.4245	0.8541**	0.3214
Shoulder width × (bicycle lane) (β_6)	−0.2075	0.1323	—	—	—	—

~Statistically significant at the 0.1 level.
*Statistically significant at the 0.05 level.
**Statistically significant at the 0.01 level.
***Statistically significant at the 0.001 level.
—No data.
AADT = annual average daily traffic.

The model estimates the expected frequency for target crashes at a site (i) and for a year (j), as shown in the equation in figure 7.

$$N_{ij} \sim Poisson(\mu_{ij})$$

Figure 7. Equation. Site-level Poisson distribution of yearly crashes.

Where:

N_{ij} = number of target crashes at site i in year j.
μ_{ij} = average yearly number of crashes at site i in year j.

The yearly expectation of crashes is further represented in the parameters shown in the equation in figure 8.

$$\mu_{ij} = RE_i \cdot Length_i^{\alpha_1} \cdot AADT_{ij}^{\alpha_2} \cdot ADBT_{ij}^{\alpha_3} \cdot \exp(\boldsymbol{X}_i' \cdot \boldsymbol{\beta})$$

Figure 8. Equation. Parameterized yearly expectation of crashes.

Where:

RE_i = baseline crash expectation at site i (estimated as a random effect).
$Length_i$ = segment length at site i.
$AADT_{ij}$ = AADT at site i and year j.
$ADBT_{ij}$ = ADBT at site i and year j.
X_i = vector of p independent variables (including bicycle lane) at site i.
α_1, α_2, α_3, β = the set of 3+p model coefficients (estimated as fixed effects across the complete dataset).

The model shows that the expected year-by-year variation is a result of yearly variations of average annual daily traffic (AADT) and ADBT.

The set of RE_i is modeled to follow the log-normal distribution with population-level parameters μ_0 and σ_0. Both parameters are subject to estimation by the model. The σ_0 parameter is estimated as the standard deviation from the estimated RE_i values in the model link scale. This parameter is a measure of unaccounted variability between sites in excess of the variability attributable to the fixed effects and can be used to estimate the amount of Poisson overdispersion present in the data.

All model variables other than length, AADT, and ADBT were included in the model in the exponential form. For clarity, the last term in figure 8 is implicit of multiple variables and can be expanded, as shown in the equation in figure 9.

$$\boldsymbol{X}' \cdot \boldsymbol{\beta} = X_1 \cdot \beta_1 + X_2 \cdot \beta_2 + \cdots + X_{\boldsymbol{p}} \cdot \beta_p$$

Figure 9. Equation. Parameterization of explanatory variables in regression model.

Where:

X_p = independent variable in the model other than length, AADT, and ADBT.

β_p = corresponding estimated coefficient.

Model results shown in table 5 indicate a consistent safety-performance link to AADT and segment length, as expected. The effect of bicycle-lane presence was consistent with expectation in two out of the three models (fatal and injury, and property-damage only (PDO)). The effect may appear counterintuitive in the model results for total crashes, but there is an interaction with shoulder width in the model that makes it difficult to judge the overall direction of the effect. Regardless, the bicycle-lane effect was found to be statistically insignificant in every case.

Researchers attempted to model bicycle crashes even though only 24 bicycle crashes were available in the database for this effort. Unfortunately, significantly large SEs and ballooning dispersion in this model indicate that no reliable estimates could be obtained for this type of crash in Washington. The next section presents CMF development from the other models documented in table 5.

CMF Development

As described in chapter 3, the research team developed a set of counterfactuals to be contrasted with the crash expectations at each location in the dataset. The corresponding CMF was then estimated as the average change in safety effectiveness between crash predictions and their corresponding counterfactuals.

The matrix of counterfactuals was defined as the design matrix with the following important modifications:

- The variable indicating no bicycle lane was changed to indicate a bicycle lane for the after period at sites without a bicycle lane in the design matrix.

- The variable indicating bicycle volume was modified to indicate a bicycle-volume percent increase for the after period at sites that do not have a bicycle lane in the design matrix.

- The variable indicating a bicycle lane was changed to indicate no bicycle lane for the before period at sites that have a bicycle lane in the design matrix.

- The variable indicating bicycle volume was modified to indicate a bicycle-volume percent decrease before bicycle-lane installation for the before period at sites that have a bicycle lane in the design matrix.

Shoulder width was a significant variable in the total crash model; therefore, an additional variant was calculated only in this case, which included the previous and the following modifications in the definition of the counterfactual matrix:

- The variable defining the shoulder width was increased by the width of the bicycle lanes for the before period at sites with bicycle lanes in the design matrix.

- The variable defining the shoulder width was decreased by the width of the bicycle lanes (or completely removed if smaller than bicycle-lane width) for the after period at sites without bicycle lanes in the design matrix.

- The models for total and PDO crashes are very close together because the majority of all crashes are PDO crashes at the studied facilities.

Total Crashes

The CMF for total crashes was estimated under the following two scenarios:

- The bicycle lane is added without modifying shoulder width (presumably reducing travel-lane width as available) (table 6).

- The addition of the bicycle lane results in reduced shoulder width (table 6).

Table 6. CMFs for bicycle-lane addition with and without shoulder reduction (total crashes).

Scenario	CMF	Standard Error (CMF)	95% Confidence Interval	
			Lower Limit	Upper Limit
Bicycle-lane addition and shoulder reduction[1]	0.7859	0.3009	0.3964	1.5583
Bicycle-lane addition and no shoulder reduction[2]	1.3065	0.4163	0.7285	2.3432

[1]Base condition: two 11.0-ft lanes, 2.0-ft shoulder, no median, urban arterial road.
[2]Base condition: two 15.0-ft lanes, 0-ft shoulder, no median, urban arterial road.

No additional computations for modified bicycle volumes were performed because the model for total crashes did not include bicycle volume as a variable; therefore, modifying bicycle volume in the counterfactuals does not affect estimated CMFs.

Both CMFs are statistically insignificant at the 5- and 10-percent levels.

Results in table 6 indicate that adding bicycle lanes without modifying shoulder width is not expected to produce a statistically significant difference in total crashes (i.e., the CMF is statistically equivalent to 1.0). Table 6 shows expected results when adding bicycle lanes and reducing shoulder width. The estimate indicates a statistically insignificant reduction in total crashes when adding bicycle lanes and reducing right shoulder width (i.e., the CMF is statistically equivalent to 1.0).

Fatal and Injury Crashes

Researchers estimated the safety effectiveness for fatal and injury crashes. In contrast to the calculation of CMFs for total crashes, no scenario was calculated where the shoulder width was compromised when adding a bicycle lane because the variable for shoulder width was not meaningful for the fatal and injury crash model. Bicycle volume was meaningful in the best-fitting crash model corresponding to fatal and injury crashes. Therefore, CMFs were estimated at the following three levels of hypothesized increase in ADBT from adding bicycle lanes:

- No change in ADBT.
- Increase of 20 percent in ADBT.
- Increase of 40 percent in ADBT.

Table 7 summarizes the results for these scenarios.

Table 7. CMFs for bicycle-lane addition by ADBT change (fatal and injury crashes).

Scenario	CMF	Standard Error (CMF)	95% Confidence Interval	
			Lower Limit	Upper Limit
Bicycle-lane addition and no change in ADBT	0.7717	0.3312	0.3635	1.6382
Bicycle-lane addition and 20-percent increase in ADBT	0.7736	0.3323	0.3642	1.6431
Bicycle-lane addition and 40-percent increase in ADBT	0.7755	0.3334	0.3649	1.6481

Note: Base condition included two lanes, no median, and urban arterial road. Contrast evaluation represented in 131 segment-periods without bicycle lanes and 48 segment-periods with bicycle lanes. All CMFs are statistically insignificant at the 5- and 10-percent levels.

The estimate in table 7 for no increase in ADBT suggests a crash reduction in fatal and injury crashes (i.e., CMF smaller than 1.0), but the result is statistically insignificant. Similarly, results are slightly larger for adding a bicycle lane and concur with a 20-percent increase in ADBT; however, results still suggest a crash reduction and remain statistically insignificant. Results remain statistically insignificant and concur with a 40-percent increase in ADBT for adding a bicycle lane. These results provide no evidence of a change in fatal and injury crashes when adding bicycle lanes and accounting for ADBT increases of 0, 20, or 40 percent.

PDO Crashes

The CMF calculation for PDO crashes takes similar features from the calculation for fatal and injury crashes because of similarities between these models. Bicycle volume was meaningful in the best-fitting crash model for PDO crashes; therefore, the research team estimated CMFs at the same three levels of hypothesized increase in ADBT as they did for fatal and injury crashes. Researchers did not anticipate an ADBT effect for PDO crashes because the CMF calculations for fatal and injury crashes indicated that the impact of increasing bicycle traffic was nearly imperceptible. The impact was imperceptible because the coefficient estimate for the impact of ADBT in crash frequency is negligible in the PDO model (i.e., two orders of magnitude smaller than the equivalent parameter estimate in the fatal and injury model). Only one CMF estimate was produced for PDO crashes. Table 8 summarizes this result.

Table 8. CMF for bicycle-lane addition regardless of ADBT increase (PDO crashes).

Scenario	CMF	Standard Error (CMF)	95% CI	
			Lower Limit	Upper Limit
Bicycle-lane addition regardless of changes in ADBT	0.8851	0.2697	0.5048	1.552

CI = confidence interval.

Note: Contrast evaluation represented in 131 segment-periods without bicycle lanes and 48 segment-periods with bicycle lanes. Both CMFs are statistically insignificant at the 5- and 10-percent levels.

The base condition for the scenario in table 8 was two lanes and no median on an urban arterial road. The estimate suggests a crash reduction (i.e., CMF smaller than 1.0). The SE is somewhat smaller than the SE for the fatal and injury CMF in table 7. Both results, however, are statistically insignificant (i.e., the CMF is statistically equivalent to 1.0).

Discussion of Results

Researchers estimated CMFs for the addition of bicycle lanes by reducing lane or shoulder width in Washington. These CMFs were estimated for three crash types: total, fatal and injury, and PDO. All CMF estimates were smaller than 1.0 except for one scenario associated with the CMF for total crashes. The corresponding SEs, however, indicated that CMFs produced using data from Washington were not statistically different from 1.0, meaning that there was no statistical evidence from the analysis that supported that the addition of bicycle lanes was linked to a shift in safety performance. The research team speculates that the small sample size is probably a reason for these statistically insignificant results, in addition to the possibility that there is no safety impact (for all roadway users) to adding bicycle lanes. For example, only 24 bicycle crashes were identified in the complete dataset, which prevented the development of a reliable model for bicycle crashes.

Texas

Like the Washington dataset, the analysis of Texas data began by exploring data to understand the underlying relationships. Given the observed ranges for key variables in the exploratory analysis, the research team applied appropriate filters to the initial dataset to ensure uniformity for analysis. Researchers anticipated that applying PS weights would even out any remaining imbalances in the data before the analyses.

In contrast with Washington data that had yearly granularity in bicycle volumes, a model-based estimate of ADBT offers limited value when predicting on a yearly basis because of the inelasticity of the predictors in the direct-demand models at the yearly level. The research team calculated an average value for the period of analysis (2015–2018) and combined other safety data for 4 yr.

Researchers anticipated that arterial roads would have different safety performance features than collectors and urban streets because of the way arterial roads operate, the amount of traffic they carry, and their geometric design. Data were divided into two groups: (1) arterials and (2) collectors and local streets. The research team further divided data based on the presence of cross sections of two and four lanes. Additional filtering was dictated by the ranges of key variables represented in the subset of sites with bicycle lanes (e.g., speed limits and AADT). Researchers discarded 878 segments in this step. The following four datasets were prepared for statistical analysis:

- Four-lane urban arterials (n = 697).
- Two-lane urban arterials (n = 188).
- Four-lane urban connectors and local streets (n = 372).
- Two-lane urban connectors and local streets (n = 1487).

The four datasets represent 2,744 segments for analysis in Texas.

For each of the four datasets, the research team developed a PS model for the presence of a bicycle lane. The research team developed overlap weights from the PS values as proposed by Li et al. (2018) and then confirmed the impact of these weights in balancing the datasets. Figure 10 through figure 13 illustrate the impact of PS weights on the balance of covariates. Figure 10 shows the distributions in the Texas dataset of the total width of lanes for four-lane segments when a bicycle lane is present.

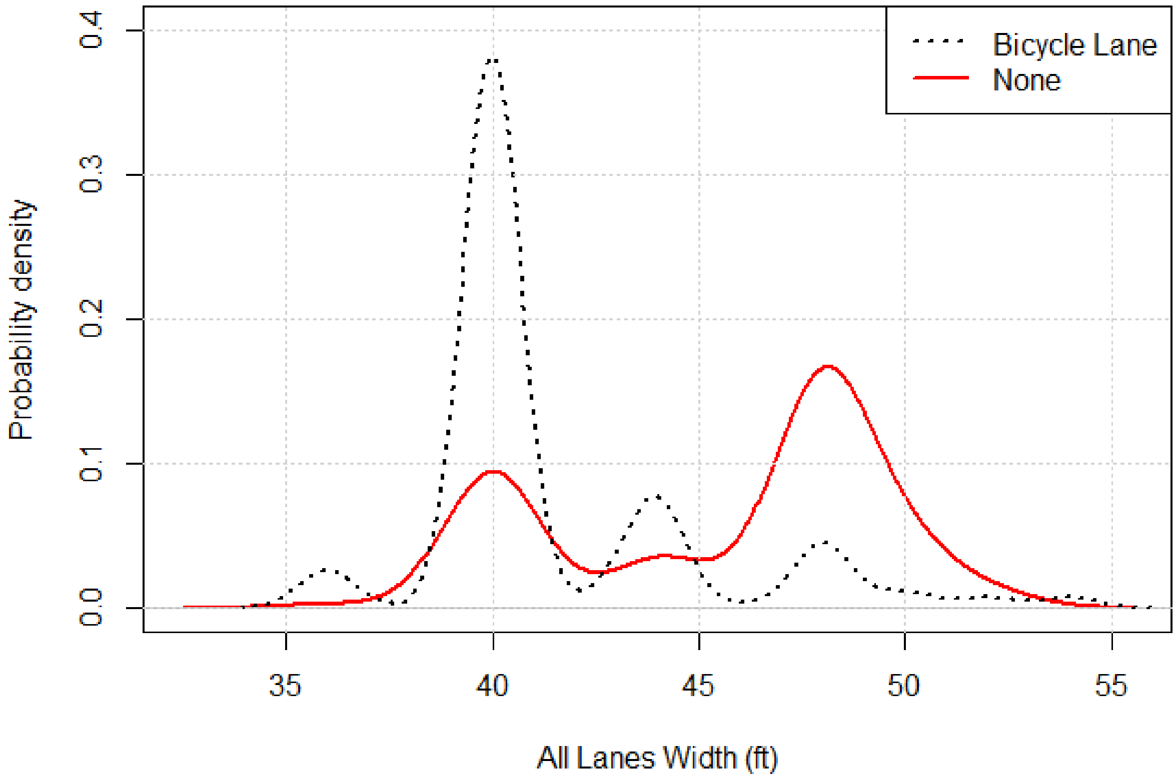

Figure 10. Graph. All-lanes width unweighted distributions by presence of bicycle lane in Texas (four-lane segments).

Figure 10 shows a clear imbalance. Among the sites with bicycle lanes, the lanes are narrower (an average lane width of 10 ft). In contrast, the largest mode in the sites without bicycle lanes is about 48 ft (or an average of 12 ft per lane). This imbalance is consistent with the expectation that sites with bicycle lanes were probably constructed while reducing the width of the previously available motor-vehicle lanes and any available shoulders. After applying PS weights, the comparison by the presence of a bicycle lane is more balanced for this covariate, as depicted in figure 11.

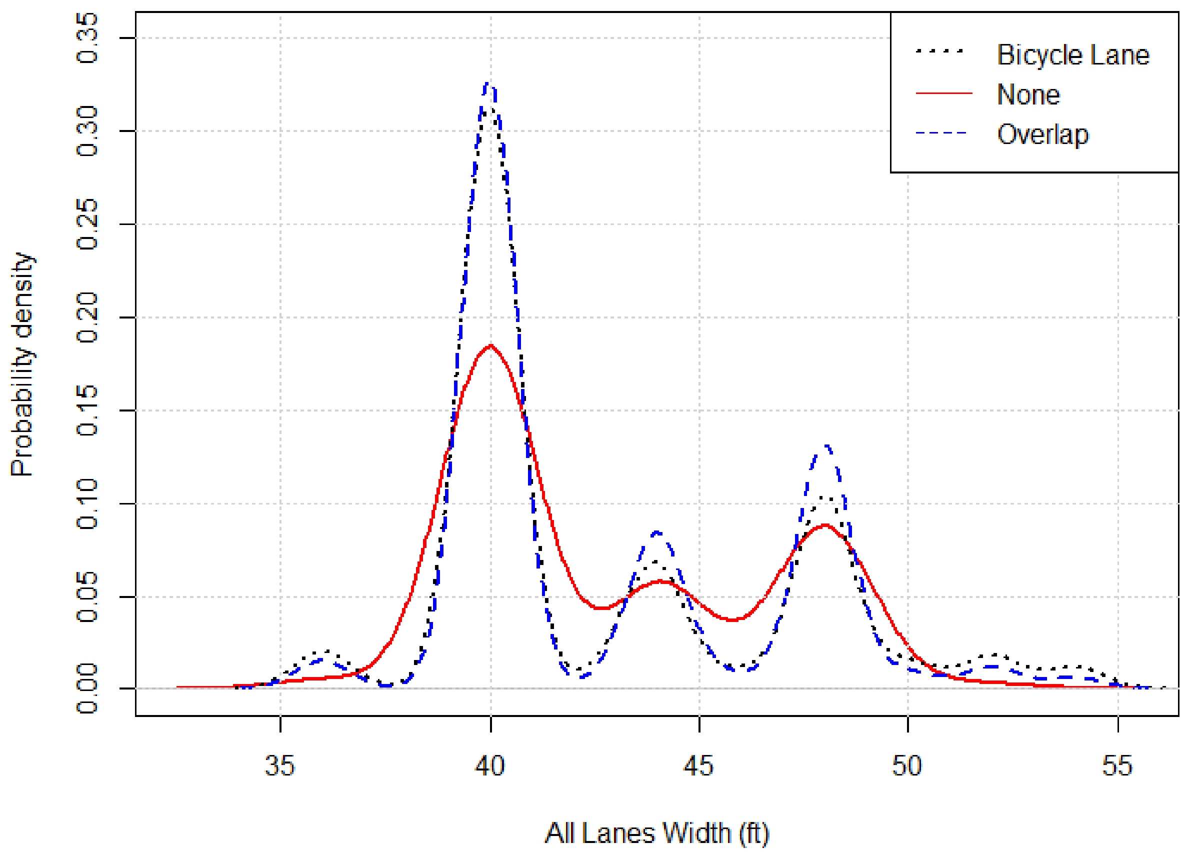

Figure 11. Graph. All-lanes width weighted distributions by presence of bicycle lane in Texas (four-lane segments).

The balancing effect of PS weights on AADT rates is shown when comparing figure 12 and figure 13.

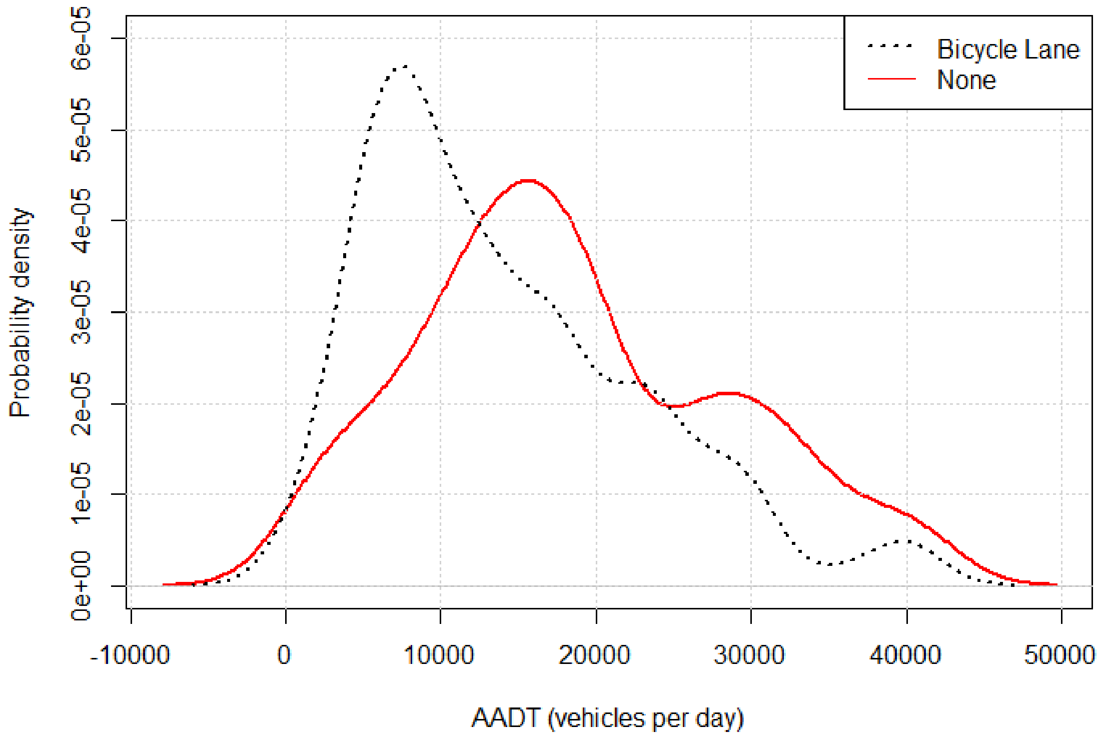

Figure 12. Graph. AADT unweighted distributions by presence of bicycle lane in Texas (four-lane segments).

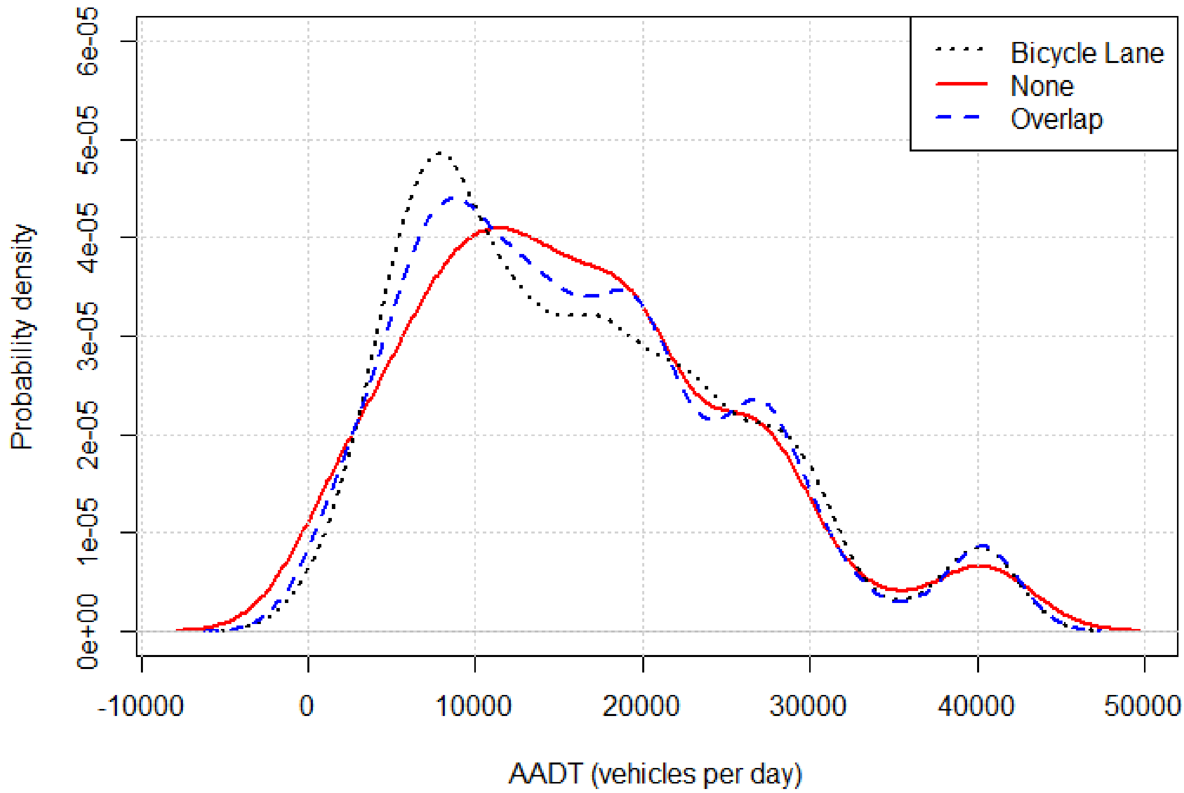

Figure 13. Graph. AADT weighted distributions by presence of bicycle lane in Texas (four-lane segments).

The AADT rates are similar in their original distributions. In the weighted distributions, AADT values larger than 50,000 vehicles per day (only available at sites without bicycle lanes) are virtually excluded by assigning near-zero weights.

Data-Analysis Results

The model for these analyses was a negative binomial GLM, implying crash observations in the dataset follow the marginal distribution of a Poisson–gamma mixture crash-generation process. This approach implies site variability as a gamma-distributed nuance parameter, whereas the crashes within a site are modeled as a Poisson variable. Therefore, any Poisson overdispersion present in the data is captured in the variability of the random effects. An estimate of Poisson overdispersion can then be constructed and reported, analogous to the dispersion parameter in the negative binomial distribution.

The general functional form of the model is such that for a site *i* and year *j*, the model estimates the expected frequency for target crashes as in figure 14.

$$N_i \sim Poisson - gamma(\mu_i)$$

Figure 14. Equation. Poisson–gamma distribution at the site level.

Where:

N_i = number of yearly target crashes at site i.

μ_i = average yearly number of crashes at site i.

The yearly expectation of crashes is further parameterized as shown in figure 15.

$$\mu_i = np_i \cdot Length^{\alpha_1} \cdot AADT^{\alpha_2} \cdot ADBT^{\alpha_3} \cdot \exp(\boldsymbol{X}' \cdot \boldsymbol{\beta})$$

Figure 15. Equation. Parameterization of crash-expectation parameter.

Where:

np_i = gamma mixture nuance parameter for crash expectation at site i (implied in the model, estimated as overall Poisson overdispersion).

Length = segment length.

AADT = AADT.

ADBT = ADBT.

X = set of p independent variables (including bicycle lane).

α_1, α_1, α_1, β = the set of 3+p model coefficients (estimated as fixed effects across the complete dataset).

In the model, the expected year-by-year variation is a result of the yearly variations of AADT and ADBT.

The set of np_i is not explicitly modeled in this case because it is in a mixed-effects model, but it is assumed to follow the gamma distribution. The estimation yields population-level parameters μ_0 and d_0. These parameters are also subject to estimation by the model. The d_0 parameter is estimated as the inversed Poisson overdispersion parameter, which captures unaccounted variability between sites in excess of the variability attributable to a heterogeneous Poisson process.

All model variables other than length, AADT, and ADBT were included in the model in the exponential form. The last term in figure 15 is implicit of multiple variables and can be expanded as shown in figure 16.

$$\boldsymbol{X}' \cdot \boldsymbol{\beta} = X_1 \cdot \beta_1 + X_2 \cdot \beta_2 + \cdots + X_{\boldsymbol{p}} \cdot \beta_p$$

Figure 16. Equation. Parameterization of explanatory variables in the model.

Where:

X_p = independent variable in the model other than length, AADT, and ADBT.

β_p = corresponding estimated coefficient.

Results from the analyses are presented in sections by facility type. Each section presents the results from the analyses that were feasible for each data subset. Table 9 through table 16 indicate statistical model coefficient estimates and CMF estimates for the addition of bicycle lanes in the different facility types.

For CMF development, researchers followed a similar but simpler approach than the approach used in Washington data. Model safety performance prediction estimates were developed and compared for the average conditions present among sites with bicycle lanes and sites without bicycle lanes. In other words, when comparing sites with and without bicycle lanes, the comparison was made while fixing the averages of other critical covariates at the group average. For example, in the CMF estimation at urban arterial roads and collectors, lane width for sites with bicycle lanes was fixed at 11 ft and for sites without bicycle lanes was fixed at 12 ft. These measurements reflect the average condition at those sites. When appropriate, scenarios were developed for multiple levels of other significant covariates in the models, such as ADBT, lane width, and shoulder width.

Safety Evaluation of Four-Lane Undivided Urban Arterials

Table 9 presents the best-fitting models for four-lane undivided urban arterials. As described previously in this chapter, statistical models were fitted to the corresponding crash severities with sufficient representation in this dataset. The model specification for this analysis was a negative binomial GLM. PS weights were applied to the regression to counter imbalances in key covariates in the data.

Table 9. Coefficient estimates for urban four-lane arterial crash-prediction models in Texas (*n* = 697).

Parameter	All Crashes		Fatal and Injury Crashes	
	Coefficient	Standard Error	Coefficient	Standard Error
Intercept (β_0)	−3.275*	1.973	−11.580	1.245
Bicycle lane (β_1) (1 = yes, 0 = no)	−0.104	0.401	0.051	0.427
AADT (α_1)	0.346	0.2198	1.000	N/A
AADT (β_2)	N/A	N/A	-6.489×10^{-5}**	1.506×10^{-5}
ADBT (β_3)	N/A	N/A	N/A	N/A
Length (α_2)	1.000	N/A	0.081	0.148
Roadbed width (ft) (β_4)	N/A	N/A	−0.0251	0.0169
Total shoulders width (ft) (β_5)	N/A	N/A	0.1240**	0.0616
Total shoulders width (ft) (no bicycle lane only) (β_6)	N/A	N/A	−0.1177**	0.0591
Speed limit (β_7)	0.060**	0.0304	0.0663*	0.0353

*Statistically significant at the 10-percent level.
**Statistically significant at the 5-percent level.
N/A = not applicable.

Analysis was not performed on bicycle-only crashes because this dataset had only four such crashes recorded for the 4-yr period of analysis.

Similar to the results from the Washington database, the modeling results in table 9 indicate consistent safety performance associations with AADT, segment length, and some cross-sectional dimensions. The coefficient indicating the effect of a bicycle lane was statistically insignificant. Table 10 shows the corresponding CMFs from these models.

Table 10. CMFs for bicycle-lane addition regardless of ADBT (four-lane undivided urban arterials).

Scenario	Crash Type or Severity	CMF	Standard Error (CMF)	95% Confidence Interval Lower Limit	95% Confidence Interval Upper Limit
Bicycle-lane addition and shoulder reduction[1]	KABCO	0.901	0.3850	0.3964	1.5583
Bicycle-lane addition and shoulder reduction[1]	KABC	1.01	0.464	0.407	2.522
Bicycle-lane addition and shoulder reduction[2]	KABC	1.032	0.416	0.4463	2.388

[1]Base condition: four 12.0-ft lanes, 6.0-ft shoulder, no median, four-lane urban arterial road.
[2]Base condition: four 12.0-ft lanes, 3.0-ft shoulder, no median, four-lane urban arterial road.
KABCO = fatal, injury, or no-injury crash.
Note: All CMFs are statistically insignificant at the 5- and 10-percent levels.

The magnitude of the results in table 10 suggests mild safety improvements from adding a bicycle lane. However, considering the SEs, these estimates are statistically insignificant (i.e., the CMF is statistically equivalent to 1.0) and therefore offer no evidence of any change in total and fatal and injury crash expectancy.

Safety Evaluation of Two-Lane Undivided Urban Arterials

Table 11 presents the best-fitting models for two-lane undivided urban arterials. As described previously in this chapter, statistical models were fitted for the corresponding crash responses with sufficient representation in this dataset. The model specification for this analysis was a negative binomial GLM. PS weights were applied to the regression to counter imbalances in key covariates in the data.

Table 11. Coefficient estimates for urban two-lane arterial crash-prediction models in Texas (*n* = 188).

Parameter	All Crashes: Coefficient	All Crashes: Standard Error
Intercept (β_0)	−4.361	3.703
Bicycle lane (β_1) (1 = yes, 0 = no)	−0.584	0.370
AADT (α_1)	0.84*	0.463
AADT (β_2)	-9.169×10^{-5}**	3.837×10^{-5}
Length (α_2)	0.03988	0.2035
Roadbed width (ft) (β_3)	−0.05333*	0.02958

*Statistically significant at the 10-percent level.
**Statistically significant at the 5-percent level.

Analysis was not performed on fatal and injury or bicycle-only crashes because this dataset had only 14 fatal and injury and 3 bicycle-only crash types for the 4-yr period of analysis.

The model results in table 11 indicate consistent safety performance associations with AADT, segment length, and the cross-sectional dimension, as expected. The coefficient indicative of the effect of a bicycle lane, however, was statistically insignificant. This can be confirmed in table 12, which shows the corresponding CMFs from this model.

Table 12. CMF for bicycle-lane addition regardless of ADBT (two-lane undivided urban arterials).

Scenario	Crash Type or Severity	CMF	Standard Error (CMF)	95% Confidence Interval Lower Limit	95% Confidence Interval Upper Limit
Bicycle-lane addition and shoulder reduction	KABCO	0.558	0.348	0.270	1.153

KABCO = fatal, injury, or no-injury crash.
Note: Base condition: two 12.0-ft lanes, 6.0-ft shoulder, no median, two-lane urban arterial road. CMFs are statistically insignificant at the 5- and 10-percent levels.

The magnitude of the result in table 12 suggests a large safety improvement when adding a bicycle lane without changing total cross-section width. However, this estimated improvement is statistically insignificant (i.e., the CMF is statistically equivalent to 1.0, considering the SE) and offers no evidence of any change in total crash expectancy.

Safety Evaluation of Four-Lane Undivided Urban Collectors and Local Roads

Table 13 presents the best-fitting models for four-lane undivided urban collectors. As with the case of urban undivided arterials, statistical models were fitted for the corresponding response variables with sufficient representation of crashes in this dataset. The model specification for these analyses was again GLMs, differentiating between negative binomial for total crashes and binomial regression for fatal and injury crashes. This differentiation was needed because only 52 fatal and injury crashes were present in this dataset. However, similarly to previous analyses, PS weights were applied to the regression to counter imbalances in key covariates in the data.

Table 13. Coefficient estimates for urban four-lane undivided urban collectors and local roads crash-prediction models in Texas (*n* = 372).

Parameter	All Crashes Coefficient	All Crashes Standard Error	Fatal and Injury Crashes Coefficient	Fatal and Injury Crashes Standard Error
Intercept (β_0)	−1.028	1.339	−7.238*	3.169
Bicycle lane (β_1) (1 = yes, 0 = no)	−0.665*	0.286	−0.634	0.453
AADT (α_1)	0.205	0.154	0.534	0.360
ADBT (β_2)	0.0053	0.0064	N/A	N/A
Length (α_2)	0.519*	0.243	1.0000	N/A
Speed limit (mph) (β_4)	N/A	N/A	0.0675*	0.0288

*Statistically significant at the 5-percent level.
N/A = not applicable.

Analysis was not performed for bicycle-only crashes because this dataset had only 26 such crashes recorded for the 4-yr period of analysis.

Results in table 13 indicate safety performance associations with segment length and presence of bicycle lane. Statistical associations were not found for individual cross-sectional dimensions as in previous analyses, perhaps because of the limited size of this dataset. The coefficient indicating the effect of bicycle lanes on total crashes was significant for total crashes but not significant for fatal and injury crashes. Table 14 shows the estimated CMFs corresponding to each crash type.

Table 14. CMFs for bicycle-lane addition regardless of ADBT (four-lane undivided urban collectors and local roads).

Scenario	Crash Type or Severity	CMF	Standard Error (CMF)	95% Confidence Interval	
				Lower Limit	Upper Limit
Bicycle-lane addition by shoulder or lane reduction (no change in ADBT)	KABCO	0.514*	0.258	0.293	0.901
Bicycle-lane addition by shoulder or lane reduction (20% increase in ADBT)	KABCO	0.542*	0.271	0.301	0.974
Bicycle-lane addition by shoulder or lane reduction (50% increase in ADBT)	KABCO	0.586	0.316	0.300	1.145
Bicycle-lane addition by shoulder or lane reduction (regardless of ADBT)	KABC	0.530	0.506	0.230	1.221

*Statistically significant at the 5-percent level.
KABCO = fatal, injury, or no-injury crash.
Note: Base condition: 11.0-ft lanes, no shoulder, no median, four-lane urban collector or local road. Initial ADBT assumed at 50 bicyclists per day (mean value in the dataset).

The result for total crashes in table 14 suggests a statistically significant reduction in total crashes when adding bicycle lanes to collectors and local roads in urban locations while experiencing up to a 20-percent increase in ADBT. This bicycle-lane addition would be achieved by reducing lane width without acquiring additional right-of-way, as the base conditions show. The magnitude of the other two CMFs in table 14 (one for total crashes and one for fatal and injury crashes) would still suggest a decrease in crash experience, but these estimates were statistically insignificant. In other words, the evaluation did not produce statistical evidence supporting that this CMF differs from a 1.0 value for total crashes when ADBT increases by 50 percent after construction of bicycle lanes. Similarly, the CMF estimate for KABC crashes was not statistically significant.

Safety Evaluation of Two-Lane Undivided Urban Collectors and Local Roads

Table 15 presents the best-fitting models for two-lane undivided urban collectors. Statistical models were fitted for the corresponding response variables with sufficient representation of crashes in this dataset. The model specification for these analyses was GLMs, the negative binomial for total crashes, and binomial regression for fatal and injury crashes (because there were only 52 fatal and injury crashes in this dataset). However, similar to previous analyses, PS weights were applied to the regression to counter data imbalances in key covariates.

Table 15. Coefficient estimates for urban two-lane undivided urban collectors crash-prediction models in Texas (n = 1487).

Parameter	All Crashes		Fatal and Injury Crashes	
	Coefficient	Standard Error	Coefficient	Standard Error
Intercept (β_0)	5.094**	0.9176	−4.400**	2.140
Bicycle lane (β_1) (1 = no, 0 = yes)	−0.8847**	0.2105	−1.763*	0.949
AADT (α_1)	−0.5573**	0.1205	0.006	0.141
AADT (β_2)	3.36×10^{-04}**	6.56×10^{-5}	N/A	N/A
Length (α_2)	0.5574**	0.1051	0.681**	0.298
ADBT (α_3)	−0.3031**	0.1398	0.462	0.521
ADBT (no bicycle-lane only) (β_3)	0.02986**	0.004782	−0.051*	0.026

*Statistically significant at the 10-percent level.
**Statistically significant at the 5-percent level.
N/A = not applicable.

Analysis was not performed for bicycle-only crashes because this dataset had only 26 such crashes recorded for the 4-yr period of analysis.

Results in table 15 indicate consistent safety performance associations with AADT, segment length, and ADBT. Statistical associations were not found for cross-sectional dimensions despite the large dataset. The two coefficients indicating the effects of a bicycle lane (e.g., a shift in the intercept and a slope by ADBT level) were significant for total crashes and marginally significant (at the 10-percent level) for fatal and injury crashes. When preparing contrasts that combine such effects for each crash type, table 16 shows that the corresponding CMFs from the total-crashes model are statistically significant, unlike CMFs for fatal and injury crashes.

Table 16. CMFs for bicycle-lane addition regardless of ADBT (two-lane undivided urban collectors and local roads).

Scenario	Crash Type or Severity	CMF	Standard Error (CMF)	95% Confidence Interval	
				Lower Limit	Upper Limit
Bicycle-lane addition by shoulder or lane reduction (no change in ADBT)	KABCO	0.734*	0.091	0.595	0.904
Bicycle-lane addition by shoulder or lane reduction (20% increase in ADBT)	KABCO	0.694*	0.096	0.557	0.865
Bicycle-lane addition by shoulder or lane reduction (50% increase in ADBT)	KABCO	0.649*	0.108	0.507	0.830
Bicycle-lane addition by shoulder or lane reduction (no change in ADBT)	KABC	1.317	0.258	0.751	2.311
Bicycle-lane addition by shoulder or lane reduction (20% increase in ADBT)	KABC	1.433	0.279	0.785	2.617
Bicycle-lane addition by shoulder or lane reduction (50% increase in ADBT)	KABC	1.589	0.343	0.776	3.255

*Statistically significant at the 5-percent level.
KABCO = fatal, injury, or no-injury crash.
Note: Base condition: 12.0-ft lanes, no shoulder, no median, two-lane urban-collector or local roads. Initial ADBT assumed at 40 bicyclists per day (mean value in the dataset).

The magnitude of the results for total crashes in table 16 suggests that adding bicycle lanes would reduce total crashes at two-lane undivided collectors and local roads in urban locations. Adding a bicycle lane would be achieved by reducing lane width without acquiring additional right-of-way, as shown in the base conditions. When considering the SEs for total crashes, estimates were found to be significantly different from 1.0 (i.e., statistical evidence of the CMF smaller than 1.0). The magnitudes of the CMFs for fatal and injury crashes suggest an increase in fatal and injury crashes, but when considering the SE of the estimates, the evaluation did not produce statistical evidence supporting that these CMFs differ from a 1.0 value (i.e., no change in fatal and injury crashes).

Discussion of Results

Researchers estimated CMFs for the addition of bicycle lanes at the expense of lane or shoulder width using the Texas database in a procedure compatible with the work documented previously in this report about estimating CMFs from the Washington database. Key differences between the two analyses include the following:

- In contrast with Washington, ADBT values were estimated from direct-demand models in Texas, which allowed for the development of a larger database.

- Data were aggregated for the 4 yr of evaluation in Texas to reflect that the ADBT estimate from the direct-demand models does not vary by year.

- The database was divided into four subsets for analysis because of the rich sample size in Texas.

CMF estimation initially considered three crash types: total, fatal and injury, and bicycle crashes. However, the number of bicycle crashes was too small in each of the four evaluation cases, which precluded the development of CMFs for this crash type. Most estimated CMFs were statistically insignificant, except for the following cases.

Total crashes in four-lane undivided collectors and local roads:

- 0.514 CMF for addition of a bicycle lane with no increase in ADBT.
- 0.542 CMF for addition of a bicycle lane with a 20-percent increase in ADBT.

Total crashes in two-lane undivided collectors and local roads:

- 0.734 CMF for addition of a bicycle lane with no increase in ADBT.
- 0.694 CMF for addition of a bicycle lane with a 20-percent increase in ADBT.
- 0.649 CMF for addition of a bicycle lane with a 50-percent increase in ADBT.

All the statistically significant CMFs indicate crash reductions. CMFs were developed for fatal and injury crashes, but none were found to be statistically significantly different from 1.0. A surprising finding was the lack of statistical significance for all CMFs developed for undivided urban arterials.

The trends observed in general, despite the statistical insignificance of the results, indicate that increasing the bicycle volume tends to impact the value of the CMF estimate, which is consistent with the expectation that ADBT plays a key role in explaining crash occurrence at the facilities under study.

SUMMARY

This chapter documented the statistical evaluations and steps taken to develop CMFs using databases from States represented in this study. Separate analyses were implemented for each dataset to develop statistical models for the expected crashes in each dataset considering differences in data structure. Using predictions from the models, the research team computed CMFs by contrasting crash predictions at each site with appropriate counterfactuals that reflect changes in site conditions from the implementation of the safety improvement under evaluation (i.e., implementing bicycle lanes by reducing lane or shoulder width). Chapter 5 outlines the B/C evaluation of the safety improvements under consideration.

CHAPTER 5. ECONOMIC ANALYSIS

The research team conducted an economic analysis to estimate B/C ratios for the addition of bicycle lanes on urban arterials, collectors, and city streets. Researchers used the Texas CMF for total crashes on collectors and local roads (CMF = 0.734), hypothesizing no increase in levels of bicycle traffic. This CMF is a conservative assumption because the increase in bicycle traffic was found to reduce the value of the CMF, as shown in the previous chapter.

To perform a B/C analysis, researchers followed the procedures recommended in the FHWA technical document titled *Highway Safety Benefit-Cost Analysis Guide* (Lawrence et al. 2018). The value of a statistical life (VSL) was obtained from the most recent memorandum on the U.S. Department of Transportation (USDOT) website (Trottenberg and Rivkin 2016). The recommended range for a VSL is from $5.2 to $12.9 million in 2012 dollars. The range for 2001 dollars allowed the research team to compute the underlying geometric rate of inflation. The range for 2016, therefore, was determined to be between $5.7 and $14.9 million. A nominal value of $10.08 million was adopted for this evaluation.

According to the Pedestrian and Bicycle Information Center, the cost of adding bicycle lanes varies depending on project details (FHWA 2015). This source argues that providing narrow lanes and reducing lane width by adding bicycle lanes costs at least $5,000 per mile, a cost that varies widely depending on pavement condition. Restriping to reduce lanes, adding bicycle lanes, or adding onstreet parking is estimated to cost approximately $5,000 to $20,000 per mile. Interventions circumscribed to the existing right-of-way (as in this evaluation) have lower associated costs (Weigand et al. 2013). Striping of travel lanes, shoulders, and bicycle lanes costs approximately $14,000 per mile for a 4-inch solid line on one side of a lane or as a trail centerline. Restriping 1 mi to add bicycle lanes and reduce the number of traffic lanes costs approximately $20,000 to $48,000 per mile, depending on the number of old lane lines to be removed. Widening the road to add lanes as part of the roadway construction project has a reconstruction cost of approximately $287,000 to $300,000 per mile (Rivers and Associates, Inc. 2016).

Litman (2017) recommends considering the direct benefits to users from improved active transport conditions and the benefits to society from increased walking and cycling activity and reduced motor-vehicle travel (resulting in more compact and multimodal community development) for B/C analysis (Litman 2017). The Pedestrian and Bicycle Information Center estimated that the cost of congestion for the average commuter was $960 per year in 2014. The cost of owning and operating a motor vehicle, however, was estimated at $8,849 per year in 2018, and the cost of biking was $308 per year.

Because adding a bicycle lane involves restriping existing rights-of-way, the team used a range of $5,000 to $20,000 to estimate the initial cost of installing a bicycle lane and a service life of 3 yr. Researchers assumed a cost of $1,000 per mile for annual maintenance costs.

COST OF CRASHES AND B/C RATIO

Researchers computed the average cost of a crash using all severities and guidance from USDOT and the distribution of severe crashes observed in each State (Trottenberg and Rivkin 2016).

Using a VSL of $10.08 million, the team estimated the average cost of a crash in Texas at $204,160. The present value of an annuity factor for a 3-yr service life was determined to be 2.94. The benefit of changing to commuting by bicycle was excluded because of the assumed zero increase in ADBT after the implementation of the bicycle lane. The B/C ratio was estimated for Texas at 16.61 after the installation of bicycle lanes.

SUMMARY

This chapter described the economic analysis performed to estimate the economic effectiveness of implementing bicycle lanes by reducing lane or shoulder width on two-lane undivided urban collectors and local streets in Texas. The chapter began by outlining the resources and assumptions involved in developing the B/C ratio. The economic evaluation yielded a B/C ratio larger than 1.0, which indicates that larger benefits than costs resulted from implementing bicycle lanes by reducing lane or shoulder width on two-lane undivided urban collectors and local streets. The following chapter presents a summary and conclusions for the project.

CHAPTER 6. SUMMARY AND CONCLUSIONS

The objective of this study was to perform a rigorous safety-effectiveness evaluation of adding a bicycle lane while reducing lane and shoulder width at urban and suburban locations that are candidates for the treatment. To accomplish the goals of this study, the research team compiled safety data from Washington and Texas. The evaluation included total, fatal and injury, PDO, and bicycle crashes. Crashes occurring at intersections were removed from analysis because the scope of the project was roadway segments.

Special emphasis was given to locations with bicycle-traffic data available because these data are an influential variable identified in past research on the safety effectiveness of the treatment. AADT is used to account for motor-vehicle exposure; likewise, bicycle traffic should reflect exposure for vulnerable bicycle users. The research team developed estimates of ADBT using actual bicycle counts for Washington and direct-demand models for Texas. In this evaluation, bicycle-traffic exposure was not only a covariate for which to control in the analysis but also an explanatory variable expected to experience a shift because of implementing a bicycle lane. A realistic assessment of the shift in safety performance (if any) after the intervention, therefore, should capture an updated safety-performance baseline (resulting from geometry modifications), as well as changes from induced bicycle traffic (increase in bicycle traffic resulting from the addition of bicycle lanes). The research team maintains that the safety effectiveness of the treatment should be estimated only by controlling for changes in exposure (i.e., for motor vehicle and bicycle) so that the estimate is comparable with currently available CMFs for other countermeasures. This study developed sets of CMF estimates for the following philosophies whenever possible: CMFs that are marginal of all exposure and CMFs under hypothesized levels of induced bicycle traffic.

The database assembled for Washington was small; therefore, the results of the analysis of Washington's data did not yield statistical evidence of any changes in safety performance from the addition of bicycle lanes. The database assembled for Texas was larger; therefore, results from Texas produced CMFs indicating statistically significant reductions for total crashes for two facility types: two-lane undivided urban collectors and local streets (0.734 statistically significant CMF for total crashes) and four-lane undivided urban collectors and local streets (0.514 statistically significant CMF for total crashes). CMFs were also estimated for fatal and injury crashes but were found to be statistically insignificant (insignificant CMF estimates for fatal and injury crashes were 1.01 for four-lane urban arterials, 0.530 for four-lane urban collectors and local roads, and between 1.31 and 1.58 for two-lane urban collectors and local roads, depending on the change in ADBT).

Regardless of the statistical insignificance of several CMFs estimated in this study, the trend suggests crash reductions at sites with bicycle lanes compared to sites without. The same trend was also true for the five CMFs found to be statistically significant. Additionally, the magnitude of the estimated CMF for four-lane collectors and local streets is sensitive to the expected increase in ADBT with bicycle lanes compared to four-lane collectors and local streets without bicycle lanes. At these facilities, the magnitude of the CMF increases with increasing ADBT, and the statistical significance of the results tends to decrease with increasing ADBT accordingly. In the case of two-lane collectors and local streets, the safety effectiveness estimate is not as sensitive to the expected increase in ADBT. The statistically significant results in this study are comparable in magnitude and direction to the CMFs developed by Park et al. (2015).

ACKNOWLEDGMENTS

The authors gratefully acknowledge the contributions from the State DOTs that provided data and valuable feedback for completion of this study. The authors specifically acknowledge the following Pooled Fund Study (PFS) and non-PFS State DOTs for providing data that were used or considered in the study: Washington and Texas.

REFERENCES

AASHTO. (1998). *Strategic Highway Safety Plan*, AASHTO, Washington, DC.

AASHTO. (1999). *Guide for Development of Bicycle Facilities.* American Association of State Highway and Transportation Officials, Washington, DC.

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

Banihashemi, M. (2016). "Effect of Horizontal Curves on Urban Arterial Crashes." *Accident Analysis and Prevention, 95*, pp. 20–26, Elsevier, Amsterdam, Netherlands.

Bonneson, J. and Pratt, M. (2009). *Roadway Safety Design Workbook*, FHWA/TX-09/0-4703-P2, Texas Transportation Institute, College Station, TX.

Booth, J. and Hobert, J. (1998). "Standard Errors of Prediction in Generalized Linear Mixed Models." *Journal of the American Statistical Association*, *93*(441), pp. 262–272, Taylor and Francis, London, UK.

Campbell, D. and Russo, M. (1999). *Social Experimentation*, Sage Publications, Thousand Oaks, CA.

Campbell, D. and Stanley, J. (1966). *Experimental and Quasi-Experimental Designs for Research.* Rand McNally, Chicago, IL.

Chen, J. and Tian, X. (2012). "Safety Analysis of Expressway's Shoulder Width." *Applied Mechanics and Materials, 256–259*, pp. 1731–1736, Trans Tech Publications, Zurich, Switzerland.

Dixon, K., Fitzpatrick, K., and Avelar, R. (2016). "Operational and Safety Trade-Offs, Reducing Freeway Lane and Shoulder Width to Permit an Additional Lane." *Transportation Research Record: Journal of the Transportation Research Board, 2588*, pp. 89–97, Transportation Research Board, Washington, DC.

FHWA. (2000). "Restriping Existing Roads with Bicycle lanes." (website) Washington, DC. Available online: https://safety.fhwa.dot.gov/PED_BIKE/univcourse/pdf/swless20.pdf, last accessed April 7, 2020.

FHWA. (2015). "Pedestrian and Bicycle Information Center." (website) Available online: http://www.pedbikeinfo.org/, last accessed March 13, 2019.

Gross, F., Jovanis, P., and Eccles, K. (2009). "Safety Effectiveness of Lane and Shoulder Width Combinations on Rural, Two-Lane, Undivided Roads." *Transportation Research Record: Journal of the Transportation Research Board, 2013*(1), pp. 42–49, Transportation Research Board, Washington, DC.

Google. (2019). Google Earth Pro v7.3.2.5779. Google LLC, Mountain View, CA.

Guo, S. and Fraser, M. (2014). *Propensity Score Analysis: Statistical Methods and Applications*, Sage Publications, Thousand Oaks, CA.

Hallett, I., Luskin, D., and Machemehl, R. (2006). *Evaluation of On-Street Bicycle Facilities Added to Existing Roadways*, FHWA/TXDOT-06/0-5157-1, Texas Department of Transportation, Austin, TX.

Hunter, W., Stewart, J., Stutts, J., Huang, H., and Pein, W. (1999). *Bicycle Lane versus Wide Curb Lanes: Operational and Safety Findings and Countermeasure Recommendations*, FHWA-RD-99-035, Federal Highway Administration, Washington, DC.

Imai, K. and Ratkovic, M. (2015). "Robust Estimation of Inverse Probability Weights for Marginal Structural Models." *Journal of the American Statistical Association, 110*(511), pp. 1013–1023.

Iowa State University. (2014). *Roadway Inventory Database (RID)*. Center for Transportation Research and Education, Ames, IA.

Johnson, R. and Wichern, D. (2007). *Applied Multivariate Statistical Analysis*, Sixth Edition. Pearson Prentice Hall, Upper Saddle River, NJ.

Jovanis, P. and Gross, F. (2007). "Estimation of Safety Effectiveness of Changes in Shoulder Width with Case Control and Cohort Methods." *Transportation Research Record: Journal of the Transportation Research Board*, *2019*(1), pp. 237–245, Transportation Research Board, Washington, DC.

Krizek, K., Gary, B., and Poindexter, G. (2006). *National Cooperative Highway Program Report 552*, *Guidelines for Analysis of Investments in Bicycle Facilities.* Transportation Research Board, Washington, DC.

Lawrence, M., Hackey, A., Bahar, G., and Gross, F. (2018). *Highway Safety Benefit-Cost Analysis Guide*, FHWA-SA-18-001. Federal Highway Administration, Washington, DC.

Lee, C., Abdel-Aty, M., Park, J., and Wang, J. (2015). "Development of Crash Modification Factors for Changing Lane Width or Roadway Segments Using Generalized Nonlinear Models." *Accident Analysis and Prevention*, *76*, pp. 83–91, Elsevier, Amsterdam, Netherlands.

Lewallen, S. and Courtright, P. (1998). "*Epidemiology in Practice: Case-Control Studies*." *Community Eye Health*, *11*(28), pp. 57–58, International Centre for Eye Health.

Li, F., Morgan, K., and Zaslavsky, A. (2018). "Balancing Covariates via Propensity Score Weighting." *Journal of the American Statistical Association*, *113*(521), pp. 390–400, Taylor and Francis, London, UK.

Li, H., Graham, D., and Majumdar, A. (2013). "The Impacts of Speed Cameras on Road Accidents: An Application of Propensity Score Matching Methods." *Accident Analysis and Prevention*, *60*, pp. 148–157, Elsevier, Amsterdam, Netherlands.

Linden, A. and Samuels, S. (2013). "Using Balance Statistics to Determine the Optimal Number of Controls in Matching Studies." *Journal of Evaluation in Clinical Practice*, *19*, pp. 968–975, International Journal of Public Health Policy and Health Services Research, Sur/Diyarbakir, Turkey.

Litman, T. (2020). *Evaluating Active Transport Benefits and Costs*. Victoria Transport Policy Institute, British Columbia, Canada. Available online: https://www.vtpi.org/nmt-tdm.pdf, last accessed October 13, 2020.

Manuel, A., El-Basyouny, K., and Islam, M. (2014). "Investigating the Safety Effects of Road Width on Urban Collector Roadways." *Safety Science*, *62*, pp. 305–311, Elsevier, Amsterdam, Netherlands.

Morrell, C., Pearson, J., and Brant, L. (1997). "Linear Transformation of Linear Mixed-Effects Models." *The American Statistician*, *51*(4), pp. 338–343, Taylor and Francis, London, UK.

NACTO. (2005). *Bike Lanes, Edge Lines, and Vehicular Lane Widths*, National Association of City Transportation Officials, New York. Available online: https://nacto.org/docs/usdg/bike_lanes_edge_lines_vehicular_lane_widths_burlington.pdf, last accessed April 7, 2020.

Olmos, A. and Govindasamy, P. (2015). "A Practical Guide for Using Propensity Score Weighting in R." *Practical Assessment, Research and Evaluation*, *20*(13), University of Massachusetts Amherst, Amherst, Massachusetts.

OpenStreetMap. (2004). Available online: https://www.openstreetmap.org/#map=5/38.007/-95.844, last accessed November 2020.

Park, J., Abdel-Aty, M., Lee, J., and Lee, C. (2015). "Developing Crash Modification Functions to Assess Safety Effects of Adding Bike Lanes for Urban Arterials with Different Roadway and Socio-economic Characteristics." *Accident Analysis and Prevention*, *74*, pp. 179–191, Elsevier, Amsterdam, Netherlands.

Pedestrian and Bicycle Information Center (2015). (website) "Narrow Lanes." National Highway Traffic Safety Administration, Washington, DC. Available online: http://guide.saferoutesinfo.org/engineering/narrow_lanes.cfm, last accessed April 2020.

Pinheiro, J. and Bates, D.M. (2000). *Mixed-Effects Models in S and S-PLUS.* Springer, New York.

Rivers and Associates, Inc. (2016). *Town of Atlantic Beach Comprehensive Bicycle Plan*, Appendix F—Cost Estimates. North Carolina Department of Transportation, Raleigh, NC. Available online: https://connect.ncdot.gov/municipalities/PlanningGrants/Documents/Atlantic%20Beach%20Bike%20Plan.pdf, last accessed April 2020.

Rosenbaum, P. and Rubin, D. (1984). "Reducing Bias in Observational Studies Using Subclassification on the Propensity Score." *Journal of the American Statistical Association*, *79*(387), pp. 516–524, Taylor and Francis, London, UK.

Sasidharan, L. and Donnell, E. (2013). "Application of Propensity Scores and Potential Outcomes to Estimate Effectiveness of Traffic Safety Countermeasures: Exploratory Analysis Using Intersection Lighting Data." *Accident Analysis and Prevention, 50*, pp. 539–553, Elsevier, Amsterdam, Netherlands.

Sasidharan, L. and Donnell, E. (2014). "Propensity Scores-Potential Outcomes Framework to Incorporate Severity Probabilities in the Highway Safety Manual Crash Prediction Algorithm." *Accident Analysis and Prevention*, *71*, pp. 183–193, Elsevier, Amsterdam, Netherlands.

Strava. (2018). *Strava Metro.* Strava, Inc. Available online: https://metro.strava.com, last accessed November 29, 2020.

Stuart, E. (2010). "Matching Methods for Causal Inference: A Review and a Look Forward." *Statistical Science: A Review Journal of the Institute of Mathematical Statistics*, *25*(1), pp. 1–21, Institute of Mathematical Statistics, Cleveland, OH.

Thomas, B., and DeRobertis, M. (2013). "The Safety of Urban Cycle Tracks: A Review of the Literature." *Accident Analysis and Prevention*, *52*, pp. 219–227, Elsevier, Amsterdam, Netherlands.

TRB. (2003). *National Cooperative Highway Research Program Report 500.* Transportation Research Board, The National Academies of Sciences, Engineering, and Medicine, Washington, DC.

TRB. (2015). *Naturalistic Driving Study: Development of the Roadway Information Database*. Transportation Research Board. Washington, DC.

TRB. (2016). *Highway Capacity Manual, A Guide for Multimodal Mobility Analysis*, Sixth Edition. Transportation Research Board, The National Academies of Sciences, Engineering, and Medicine, Washington, DC.

Trottenberg, P., and Rivkin, R. (2016). *Guidance on Treatment of the Economic Value of a Statistical Life in U.S. Department of Transportation Analyses.* United States Department of Transportation, Washington, DC. Available online: https://www.transportation.gov/regulations/guidance-treatment-economic-value-statistical-life-us-department-transportation-analyses, last accessed April 2020.

Texas A&M Transportation Institute (2019). “Texas Bicycle and Pedestrian Locations Database.” Texas A&M University, College Station, TX. Available online: https://tableau.tamu.edu/t/TTI/views/TxDOTBikePedDatabaseDisplayMasterv2/StateView?iframeSizedToWindow=true&:embed=y&:showAppBanner=false&:display_count=no&:showVizHome=no, last accessed November 2020.

TxDOT. (2020). “Open Data Portal.” (website) Texas Department of Transportation, Austin, TX. Available online: http://gis-txdot.opendata.arcgis.com/, last accessed November 2020.

Turner, S., Benz, R., Hudson, J., Griffin, G., Lasley, P., Dadashova, B., and Das, S. (2019). *Improving the Amount and Availability of Pedestrian and Bicyclist Count Data in Texas*, Texas Department of Transportation, Austin, TX.

U.S. Census Bureau. (2018). *American Community Survey*. United States Census Bureau. Washington, DC.

van Schalkwyk, I., and Washington, S. (2008). *Cost Effective Safety Improvements on Two-Lane Rural State Roads in Washington State.* Washington State Department of Transportation, Olympia, WA.

Vermeulen, K., and Vansteelandt, S. (2015). “Bias-Reduced Doubly Robust Estimation.” *Journal of the American Statistical Association 110*(511) pp. 1024–1036, Taylor and Francis, London, UK.

Wackerly, D.D., Mendenhall III, W., and Scheaffer, R.L. (2008). *Mathematical Statistics with Applications.* Seventh Edition. Thomson Brooks Cole, Cengage, Toronto, Canada.

Weigand, L., McNeil, N., and Dill, J. (2013). “Cost Analysis of Bicycle Facilities: Cases from Cities in the Portland, Oregon, Region.” *Active Living Research*, Robert Wood Johnson Foundation, Portland, OR.

RECOMMENDED CITATION: FEDERAL HIGHWAY ADMINISTRATION,
DEVELOPING CRASH MODIFICATION FACTORS FOR BICYCLE-LANE ADDITIONS
WHILE REDUCING LANE AND SHOULDER WIDTHS (WASHINGTON, DC: 2021)
HTTPS://DOI.ORG/10.21949/1521657

HRDS-20/07-21(300)E