# ROADWAY SAFETY INSTITUTE

Human-centered solutions to advanced roadway safety

# Pedestrian and Bicycle Crash Risk and Equity: Implications for Street Improvement Projects


**Greg Lindsey**
**Tao Tao**
**Jueyu Wang**
**Jason Cao**

Humphrey School of Public Affairs
University of Minnesota


Final Report

REGION 5

CTS 19-15

University of Minnesota Driven to Discover
The University of Akron
Illinois University of Illinois at Urbana-Champaign
Southern Illinois University Edwardsville
Western Michigan University

## Technical Report Documentation Page

| 1. Report No. | 2. | 3. Recipients Accession No. |
|---|---|---|
| CTS 19-15 | | |

| | |
|---|---|
| 4. Title and Subtitle<br>Pedestrian and Bicycle Crash Risk and Equity: Implications for Street Improvement Projects | 5. Report Date<br>June 2019<br>6. |
| 7. Author(s)<br>Greg Lindsey, Tao Tao, Jueyu Wang, Jason Cao | 8. Performing Organization Report No. |
| 9. Performing Organization Name and Address<br>Humphrey School of Public Affairs<br>University of Minnesota, Twin Cities<br>301 S 19th St, Minneapolis, MN 55455 | 10. Project/Task/Work Unit No.<br>CTS #2018057<br>11. Contract (C) or Grant (G) No.<br>DTRT13-G-UTC35 |
| 12. Sponsoring Organization Name and Address<br>Roadway Safety Institute<br>Center for Transportation Studies<br>University of Minnesota<br>University Office Plaza, Suite 440<br>2221 University Ave. SE<br>Minneapolis, MN 55414 | 13. Type of Report and Period Covered<br>Final Report<br>14. Sponsoring Agency Code |

15. Supplementary Notes
http://www.roadwaysafety.umn.edu/publications/

16. Abstract (Limit: 250 words)
Transportation managers need information about crash risk and equity to prioritize investments in street networks. This case study uses data from Minneapolis, Minnesota, to illustrate how estimates of pedestrian and bicycle crash risk and assessments of inequities in the distribution of that risk can inform prioritization of street improvement projects. Crash numbers and frequencies for pedestrian and bicycle crashes at intersections and mid-blocks in Minneapolis are determined for the 2005–2017 period. New models of pedestrian and bicycle crash risk at both intersections and mid-blocks that control for exposure are introduced and used to predict crashes at all intersections and mid-blocks in the city. Statistical tests are used to assess the equity of distribution of estimated crash risk between areas of concentrated poverty with majority-minority populations and other areas in the city. Crash indexes based on predicted crashes are used to illustrate how increased emphases can be placed on pedestrian and bicycle safety in street improvement rankings. Results show that pedestrian and bicycle crash risk is correlated with exposure, that different factors affect crash risk at intersections and mid-blocks, and that these factors differ for pedestrian and bicycle crashes. Results also show that mean crash risk is higher in neighborhoods with lower incomes and majority-minority populations. For street improvement projects in the city, different rankings result when segments are ranked according to modeled pedestrian and bicycle crash risk in addition to total crash rates based on historical numbers of crashes at particular locations. Results generally affirm efforts by the Minneapolis Department of Public Works to increase emphases on pedestrian and bicycle safety and equity in its prioritization of street improvements.

| 17. Document Analysis/Descriptors | 18. Availability Statement |
|---|---|
| Bicycles, Pedestrian, Safety, Data collection, Equity (Justice), Crash risk forecasting, Statistical analysis, Pedestrian safety, City planning, Traffic safety | No restrictions. Document available from: National Technical Information Services, Alexandria, Virginia 22312 |

| 19. Security Class (this report) | 20. Security Class (this page) | 21. No. of Pages | 22. Price |
|---|---|---|---|
| Unclassified | Unclassified | 53 | |

# Pedestrian and Bicycle Crash Risk and Equity: Implications for Street Improvement Projects

## FINAL REPORT

*Prepared by:*

Greg Lindsey
Tao Tao
Jueyu Wang
Jason Cao
Humphrey School of Public Affairs
University of Minnesota

## June 2019

*Published by:*

Roadway Safety Institute
Center for Transportation Studies
University of Minnesota
University Office Plaza, Suite 440
2221 University Ave. SE
Minneapolis, MN 55455

# ACKNOWLEDGMENTS

The funding for this project was provided by the United States Department of Transportation's Office of the Assistant Secretary for Research and Technology for the Roadway Safety Institute, the University Transportation Center for USDOT Region 5 under the Moving Ahead for Progress in the 21st Century Act (MAP-21) federal transportation bill passed in 2012.

The research team would like to acknowledge:

- Dr. Max Donath, director of the Roadway Safety Institute, for his vision, leadership, and commitment to increasing safety in all forms of transportation;
- Ms. Elizabeth Andrews, research coordinator at the Center for Transportation Studies, for her guidance and administrative support throughout the project; and
- Mr. Brette Hjelle and Ms. Liz Heyman with the city of Minneapolis, for their willingness to meet, discuss this project, and share data from the city.

# TABLE OF CONTENTS

## LIST OF FIGURES

# LIST OF TABLES

# EXECUTIVE SUMMARY

Transportation planners and managers need information about crash risk to prioritize investments in street networks and address community priorities related to the development of efficient and equitable transportation systems. This case study uses data from Minneapolis, Minnesota, to illustrate how estimates of pedestrian and bicycle crash risk and assessments of inequities in distribution of that risk can inform prioritization of street improvement projects. Crash numbers and frequencies for pedestrian and bicycle crashes at intersections and mid-blocks in Minneapolis are determined for the 2005–2017 period. We estimate new models of pedestrian and bicycle crash risk at intersections and mid-blocks in Minneapolis that control for vehicular, pedestrian, and bicycle exposure, use these models to predict crashes at all intersections and mid-blocks in the city, assess the equity of distribution of crash risk, show that crash risk is higher in neighborhoods with lower household incomes and higher populations of minorities, develop new indices of crash risk, and illustrate how findings and results can be used to inform ranking and prioritization of street improvement projects.

Our approach is illustrative of "systemic" approaches used in the study of roadway safety and assessment of risk described by the FHWA (Federal Highway Administration, 2018a, 2018b) and others (Carlson et al., 2018; Lindsey et al., 2018). A distinctive feature of these crash models is that they provide a measure for estimating risk at each intersection or mid-block regardless of whether crashes have occurred at the location historically. A second distinctive feature is that the models incorporate measures of vehicular, pedestrian, and bicyclist exposure.

Results show that pedestrian and bicycle crash risk at intersections and mid-blocks generally is correlated with exposure and that correlates of crash risk for different modes at intersections and mid-blocks differ. These results confirm the value of disaggregate analyses in prioritizing investments in improvements to increase safety of street networks. Results also confirm that pedestrian and bicycle crash risk is distributed unevenly throughout the city, with higher risk at intersections in lower-income neighborhoods with majority-minority populations. These differences are magnified when the central business district is excluded from analyses, indicating even greater disparities among neighborhoods. These results affirm the importance of efforts by the Minneapolis Department of Public Works to prioritize equity in ranking of street improvement projects. Finally, results show that different rankings result when network segments in the city are ranked according to modeled pedestrian and bicycle crash risk rather than total crash rates based on historical numbers of crashes at particular locations. This result confirms there are viable strategies for increasing weight given to improvements that facilitate walking and bicycling.

The study has several limitations that can be addressed over time as more data become available and additional research is undertaken. The crash models are based on a relatively small dataset, and the time periods for which the dependent and independent variables are measured are not consistent. Specifically, the years for the crash dataset and the pedestrian and bicycle counts used as measures of exposure are different, and the bicycle and pedestrian measures of exposure are two-hour, peak-hour counts, not year-round, 24-hour counts. Another limitation is that the measures of pedestrian exposure used in pedestrian mid-block crash models are estimates of pedestrians on sidewalks traveling parallel to streets, not actual mid-block pedestrian crossings. The new crash indices developed to inform project ranking could be modified or used to rank projects in many different ways. Simulation studies would be useful to assess how robust these indexes are. This study was limited to technical analyses. Future studies that involve community collaboration in developing and applying new measures of crash risk could inform future efforts to prioritize street improvements (City of Minneapolis, 2017a).

# CHAPTER 1: INTRODUCTION

Local transportation planners and policy makers need information about variation in crash risk across street networks to prioritize projects, inform investments to increase traffic safety, and address issues of equity in communities. Using Minneapolis, Minnesota, as a case study, this report illustrates how estimates of pedestrian and bicyclist exposure to risk, estimates of crash risk, and analyses of inequities in the distribution of crash risk can inform prioritization of street improvements. Crash numbers and frequencies for pedestrian and bicycle crashes at intersections and mid-blocks in Minneapolis are determined for the 2005–2017 period. Our approach is illustrative of the "systemic" approaches used in the study of roadway safety and assessment of risk described by the FHWA (Federal Highway Administration, 2018a, 2018b). This report extends recent work supported by the Roadway Safety Institute (RSI) on methods for assessing exposure to risk (Hankey & Lindsey, 2016) and using pedestrian and bicycle counts to control for exposure when estimating crash risk (Lindsey et al., 2018). This report also contributes to streams of research that document inequities in the distribution of pedestrian and bicyclist collisions (Cottrill & Thakuriah, 2010; Loukaitou-Sideris et al., 2007; Siddiqui et al., 2014).

Minneapolis has begun implementation of its 20-Year Streets Funding Plan (City of Minneapolis, 2017a). As part of its efforts to rank and prioritize potential projects, the Minneapolis Department of Public Works (DPW) conducted an outreach campaign to identify concerns of residents with its approach to project prioritization and implementation. The DPW identified the need to revise its approach to project ranking, specifically to increase emphases on pedestrian and bicycle facilities, place more "weight" on "high-pedestrian, bicycle, and transit volume streets," and integrate equity considerations (specifically for "non-white majority and low-income population(s)") (City of Minneapolis, 2017a). This research was designed to inform the city's ongoing efforts to address issues raised by residents. Specifically, the research was designed to illustrate how estimates of pedestrian and bicyclist exposure to risk can be used to model crash risk and how measures of crash risk can be used to address issues of equity.

Chapter 2 of this report is a brief literature review that focuses on approaches to modeling crash risk and studies of crash risk and equity. Chapter 3 presents our data and methods, including our models of exposure to risk, our models of crash risk, and the methods we use to assess the equity of the distribution of crash risk. Chapter 4 presents our results. These results include pedestrian and bicyclist numeric crash indices for each intersection and each mid-block in Minneapolis and the statistical tests that confirm the existence of spatial inequities in the distribution of crash risk in Minneapolis. To assess how these new measures of pedestrian and bicyclist crash risk would affect project ranking, we compare these measures with crash rates used by the city in its project ranking system. In Chapter 5, we discuss our major findings, their implications and limitations, and the need for additional research. Overall, our research confirms the importance of incorporation of estimates of exposure to risk in models of crash risk. Our research confirms that the spatial distribution of crash risk is inequitable. On average, after controlling for exposure and other factors, including the built environment and presence of traffic controls, crash risk at intersections is significantly higher in neighborhoods with higher concentrations of minority individuals and households in poverty. These findings underscore the importance of the work by the city to increase emphases on pedestrian and bicyclist traffic and to address issues of equity.

# CHAPTER 2: LITERATURE REVIEW

The principal objectives of this study are to illustrate how estimates of crash risk and assessments of inequities in the distribution of crash risk can inform prioritization of street improvement projects. This brief literature review focuses on models of crash risk and assessments of the equity of distribution of crash risk. Specifically, we summarize studies that have used regression modeling to explain pedestrian or bicycle crash rates or predict pedestrian and bicycle crashes (Section 2.1) and planning-level studies that have explored the equity of distribution of crash risk (Section 2.2).

Pedestrian crashes are collisions between a pedestrian and a motor vehicle or between a pedestrian and a bicycle. Similarly, bicycle crashes are collisions between a bicycle and a motor vehicle or between a bicycle and a pedestrian or another bicycle. Most crash datasets are incomplete and underestimate the prevalence of pedestrian and bicycle crashes because they are based on public safety records and many crashes involving pedestrians and bicyclists are not reported to the police. For example, a crash between a cyclist and a pedestrian at an intersection may not be reported unless injuries are involved.

Crash risk has been defined in many ways. Federal Highway Administration (2018a) recently has provided general guidance for assessing risk to pedestrians and bicyclists. The FHWA (2018a), defines crash risk as the likelihood that a crash will occur given exposure to risk (i.e., taking into consideration the volumes of pedestrians, vehicles, and bicycles that are present). The FHWA defines exposure as “the number of potential opportunities for a crash to occur” (Federal Highway Administration, 2018a) and notes exposure often is measured as the volumes of vehicles, pedestrians, and cyclists present with the potential to be engaged in a collision. In practice, comprehensive measures of pedestrian and bicyclist exposure rarely are available, especially at the disaggregated level, so approximations or partial estimates often are used.

Crash rates are the ratios between crash numbers and measures of exposure (i.e., traffic volumes). Crash frequency, which refers to the number of collisions at a location or in an area within a specified period of time, is used frequently in studies of crashes but technically is not a measure of risk because it does not control for exposure. Some of the studies summarized in this review analyzed crash rates and frequencies but technically not risk. We include these studies because factors associated with crash risk crash rates, and crash frequency are expected to be similar.

## 2.1 CRASH RISK AND CORRELATES OF CRASH RISK

A number of studies have investigated the impacts of different factors on crash risk and other outcomes (Table 2.1). These studies have been undertaken at different scales, ranging from area-wide analyses to disaggregated, facility-specific investigations. Area-wide analyses, or macro-level studies, often are undertaken to compare places or locations or to assess trends (Chen, 2015). Examples of units of analysis used in area-wide studies include census tracts, census block groups, or Travel Analysis Zones (Dumbaugh and Li, 2011; Siddiqui et al., 2012; Wei and Lovegrove, 2012). Examples of units of analysis in disaggregate studies include intersections (Daniels et al., 2009; Schepers et al., 2011) or corridors (Siddiqui et al., 2012; Strauss et al., 2013). The dependent variable in crash studies that use some types of regression modeling typically is some measure or index of crash frequency (Siddiqui et al., 2012), crash probability, or injury severity (Kim et al., 2007).

Correlates of crash risk can be grouped within a relatively small set of categories. These categories include exposure, built environment, traffic facilities, road characteristics, sociodemographic characteristics, and other spatial variables (Table 2.1). Measures of exposure include common, easily available measures of vehicular traffic volume such as annual average daily traffic (AADT) or comparable or related measures of pedestrian and bicycle traffic such as peak hour traffic. Researchers have found that AADT is positively associated with crash risk (Cottrill and Thakuriah, 2010; El-Basyouny and Sayed, 2013; Loukaitou-Sideris et al., 2007; National Academies of Science Engineering and Medicine, 2008; Nordback et al., 2014; Park et al., 2015; Schepers et al., 2011; Turner et al., 2011; Yasmin and Eluru, 2016). While many studies have explored the relationship between AADT and some measure of crash risk, fewer studies have tested relationships between pedestrian or bicycle counts and crash risk, mainly due to the lack of available data. The studies that have been completed indicate both pedestrian counts (National Academies of Science Engineering and Medicine, 2008; Thomas et al., 2017; Yasmin and Eluru, 2016), and bicycle counts (Nordback et al., 2014; Schepers et al., 2011; Turner et al., 2011; Yasmin and Eluru, 2016) are positively associated with crash risk (Table 2.1).

While the relationships between crashes and measures of exposure to risk have been found to be positive, some research indicates that this relationship may be non-linear, and there is debate in the literature over what has come to be known as the “safety in numbers” phenomenon (Carlson et al., 2018; Elvik, 2013, 2009; Jacobsen, 2003). The basic idea is that when pedestrian or bike traffic increases, the numbers of crashes do not increase proportionately, and the crash rate becomes smaller, indicating lower risk to individuals. Jacobsen (2003) found that the probability which a motorist collides with a pedestrian or bicyclist becomes smaller when more people walk or ride bikes. In an analysis of studies that have shown non-linearities in the risk of injury to pedestrians and bicyclists given volumes, Elvik (2009) concluded that there could be a reduction in total number of accidents if motorists shifted to walking or bicyling and that increases in non-motorized traffic volumes will not necessarily lead to increased numbers of crashes. Elvik (2013) subsequently showed that both a “safety-in-numbers” and a “hazard-in-numbers” could exist in a single crash dataset and that the researchers should differentiate between “partial” and “complete” safety-in-numbers analyses. More recently, in a systematic review and meta-analysis, Elvik and Bjørnskau (2017) concluded there is evidence of the safety-in-numbers effects for vehicles, pedestrians, and bicyclists but that causal mechanisms are not well understood. Carlson et al. (2018) concluded that the “safety in numbers” phenomenon exists in Minneaplis.

Built environment variables include measure such as population density, job density, and various categories of land use. Population and/or population density is positively associated with crash risk (Dumbaugh and Li, 2011; Gladhill and Monsere, 2012; Loukaitou-Sideris et al., 2007). Job or employment density also has been found to be positively associated with crash risk (Loukaitou-Sideris et al., 2007). Examples of measures land use include land use entropy and percentage or proportion of a specific land use such as commercial, retail, or open space. While land use entropy has been found to have a positive relationship with crash risk (Chen, 2015), other, different types of land use may have different impacts on crash risk. Commercial or business land use been shown to have a positive association with crash risk

(Cottrill and Thakuriah, 2010; Dumbaugh and Li, 2011; Gladhill and Monsere, 2012; Loukaitou-Sideris et al., 2007; Schneider et al., 2010; Thomas et al., 2017; Ukkusuri et al., 2012), while industrial land use may either a have negative association (Loukaitou-Sideris et al., 2007) or a positive one (Ukkusuri et al., 2012). Intersection density appears to be positively correlated with crash risk (Dumbaugh and Li, 2011; Gladhill and Monsere, 2012; Ukkusuri et al., 2012). Transit-related variables such as transit ridership and transit stop or station number also appear to be positively correlated with crash risk (Gladhill and Monsere, 2012; Ukkusuri et al., 2012).

Examples of traffic facilities include bike lanes (e.g., striped, buffered, or colored bike lanes), traffic signals, lighting, and measures of road geometry such as raised median (or island), road width or the number of lanes. The presence of bike lanes appears to be negatively associated with crash risk (Park et al., 2015; Schepers et al., 2011; Turner et al., 2011). Traffic signals have a mixed relationship with crash risk: while the density of traffic signals is positively associated with crash risk(Chen, 2015), the presence of a beacon may reduce the crash risk (Zegeer et al., 2017). Improving lighting fixtures at intersections or along roadway sections may reduce crash risk (Lee and Abdel-Aty, 2005; Zegeer et al., 2017). Raised medians or island are associated with lower crash risk (Schneider et al., 2010; Zegeer et al., 2017). Wider road width and more road lanes also have been found to be positively associated with crash risk (Park et al., 2015; Schepers et al., 2011; Schneider et al., 2010; Ukkusuri et al., 2012). In addition, primary roads have been shown to be positively associated with crash risk (Ukkusuri et al., 2012).

Socio-demographic characteristics include income levels, measures of numbers or density of children in nearby areas, and percentage of people of different ages, races, or ethnicities. The correlation between median household income and crash risk has differed across studies(Thomas et al., 2017; Yasmin and Eluru, 2016). Higher percentages of children in nearby communities are positively correlated with crash risk (Cottrill and Thakuriah, 2010; Schneider et al., 2010). Also, higher percentages of people of Hispanic ethnicity have been shown to have a positive relationship with crash risk (Loukaitou-Sideris et al., 2007).

Policy variables include measures related to parking and speed limits. Long-term parking costs have a positive impact on crash risk (Siddiqui et al., 2012). Higher speed limits also are positively associated with crash risk (Chen, 2015).

**Table 2.1 Selected studies on pedestrian and bicycle crashes**

| Author | Mode | | Analysis Units | | Dependent Variables | Exposure | | | Built environment | Traffic facilities | Socioeconomic | Policy | Others | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Pedestrian | Bike | Area-wide | Site-specific | | Pedestrian | Bike | Vehicle | | | | | VMT | Average speed | Bar |
| Lee and Abdel-Aty (2005) | √ | | | √ | Number of crashes | | | | | √ | √ | | | | |
| Loukaitou-Sideris et al. (2007) | √ | | √ | | Crash rate | √ | √ | √ | √ | | √ | | | | |
| National Academies of Science Engineering and Medicine (2008) | √ | | | √ | Number of crashes | √ | | √ | √ | √ | √ | | | | |
| Daniels et al. (2009) | | √ | | √ | Number of crashes | | | | | √ | | √ | | | |
| Cottrill and Thakuriah (2010) | √ | | √ | | Crash rate | | | √ | √ | | √ | | | | |
| Schneider et al. (2010) | √ | | | √ | Crash rate | √ | | √ | √ | √ | √ | | | | |
| Dumbaugh and Li (2011) | √ | √ | √ | | Number of crashes | | | | √ | | | | √ | | |
| Turner et al. (2011) | | √ | | √ | Number of crashes | | v | √ | | √ | | | | | |
| Scheper et al, (2011) | | √ | | √ | Number of crashes | | | | √ | √ | | | | | |
| Siddiqui et al. (2012) | √ | √ | √ | | Crash rate | | | | √ | √ | √ | √ | | | |
| Ukkusuri et al. (2012) | √ | | √ | | Crash rate | | | | √ | √ | √ | | | | |
| Gladhill and Monsere (2012) | √ | √ | √ | | Number of crashes | | | √ | √ | | √ | | √ | √ | |
| El-Basyouny and Sayed (2013) | √ | √ | | √ | Number of crashes | | | √ | | | | | | | |
| Nordback et al. (2014) | | √ | | √ | Number of crashes | | √ | √ | | | | | | | |
| Park et al. (2015) | | √ | | √ | Number of crashes | | | √ | √ | √ | √ | | | | |
| Chen (2015) | | √ | √ | | Number of crashes | | √ | √ | √ | √ | √ | √ | | | |
| Yasmin and Eluru (2016) | | √ | √ | | Crash rate | √ | √ | √ | √ | √ | √ | | | | √ |
| Zegeer et al. (2017) | √ | | | √ | Number of crashes | √ | | √ | | √ | | | | | |
| Thomas et al. (2017) | √ | | | √ | Number of crashes | √ | √ | | √ | √ | √ | | | | |

Other factors such as vehicle mile traveled (VMT), number of road trips, average speed, and the location of bars or drinking establishments have been shown to be correlated with crash risk. VMT and number of road trips have a positive correlation with crash risk (Chen, 2015; Dumbaugh and Li, 2011; Gladhill and Monsere, 2012). However, average speed is negatively associated with crash risk (Gladhill and Monsere, 2012). One interesting result is that the number of bars is positively associated with crash risk (Yasmin and Eluru, 2016).

These studies have greatly increased understanding of factors that influence the risk of crashes, but many have been limited because of the lack of data, particularly measures of exposure to risk. Use of exposure measures computed for large areas at the macro level (e.g., for cities (Robinson, 2005)) limits the specificity of conclusions. Many disaggregate studies have been limited to a few intersections or segments and not entire networks because of the lack of availability of counts. For example, Jonsson's (2005) scope was limited to locations (i.e., intersections, segments) where manual counts were available). Additional research to develop methods for quantifying exposure and determining its relation to crash probabilities is warranted. Among other needs, researchers and practitioners need additional tools to estimate bicycle and pedestrian traffic volumes.

## 2.2 CRASH RISK AND EQUITY

Fewer scholars have explored the relationship between crash risk and equity. Nantulya and Reich (2003) concluded that poorer groups experience a "disproportionate" burden of road traffic injuries. Siddiqui et al. (2014) suggested that the areas with low-income and minority populations require "specific attention in the investigation of pedestrian-related crash factors." Loukaitou-Sideris et al. (2007) found a statistically significant relationship between pedestrian crashes and the percentage of Latino population when they constructed models to explore the correlation between pedestrian crashes and socio-demographic variables in Los Angeles. Cottrill and Thakuriah (2010) reported that pedestrian-vehicle crashes in Chicago increased in the areas with more lower-income and minority populations.

In addition to these peer-reviewed research findings, studies in the grey literature report suggest that poorer people experience a disproportionately high pedestrian crash risk (Maciag, 2014). Most relevant to this research, a study by the city of Minneapolis has found that people living in areas with a majority of non-white and lower-income individuals experience higher crash rates than those in other areas (City of Minneapolis, 2017b).

The findings reported in equity-focused studies tend to be consistent: people in neighborhoods that are poorer and predominantly minority experience higher crash risk. However, the number of studies is relatively small, and several of the studies have been limited by the lack of availability of data, especially measures of exposure to risk. Additional studies of the relationships between crash risk and equity are warranted.

# CHAPTER 3: APPROACH, DATA, AND METHODS

In this chapter, we describe the approach, data, and methods used in our analyses. Our approach included four basic steps (Figure 3.1):

- Assembly of data;
- Estimation of pedestrian and bicyclist crash risk for intersections and mid-blocks;
- Assessment of equity of distribution of crash risk; and
- Assessment of implications for street improvement project rankings.

The data include pedestrian and bicycle crash data (Section 3.1), exposure data (i.e., measures of vehicular, pedestrian, and bicycle traffic volumes; Section 3.2), and data used as independent variables in modeling crash risk (Section 3.3). We use negative binomial regression modeling to estimate pedestrian and bicycle crash risk for each intersection and each mid-block in Minneapolis (Section 3.4). We use both standard t-tests and Lorenz curves and Gini Coefficients to assess the equity of distribution of crash risk (Section 3.5). To assess the implications of our estimates of crash risk for project ranking, we illustrate how our measures of crash risk can complement crash rates currently used by the Minneapolis DPW in its ranking system (Section 3.6).

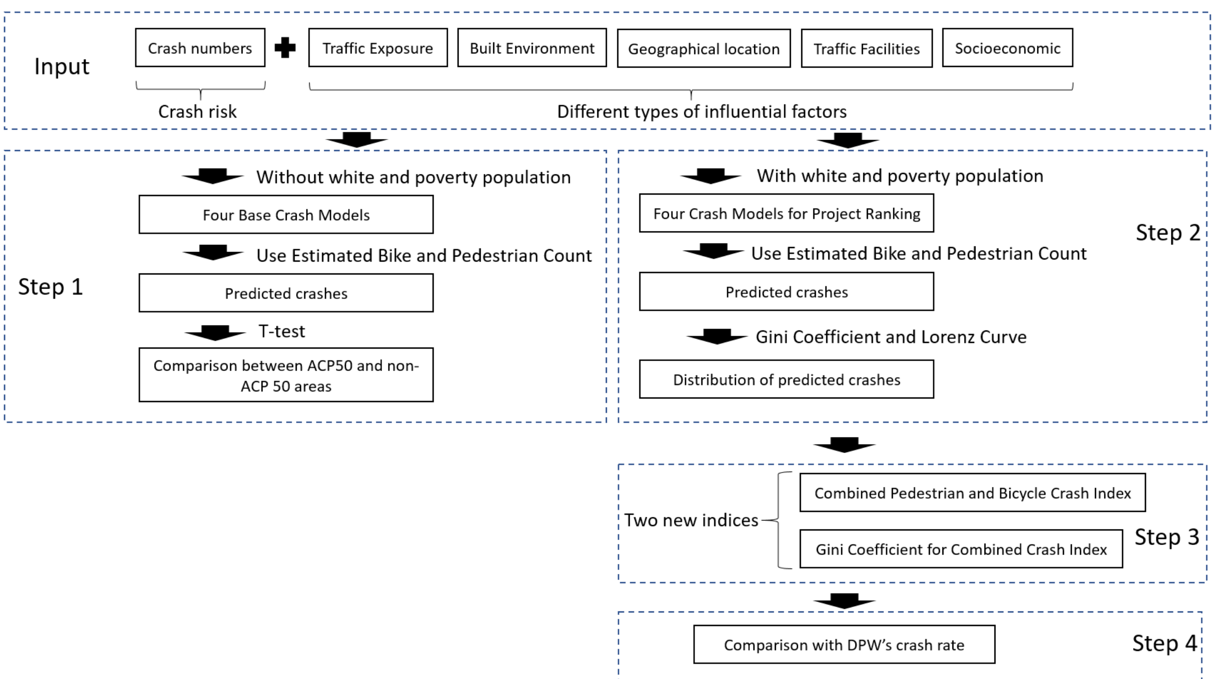



**Figure 3.1 Working flow map of the method**

## 3.1 PEDESTRIAN AND BICYCLE CRASH DATA

We define crash risk as the number (or frequency) of pedestrian and bicycle crashes that occurred at intersections and mid-blocks for the time period 2005 – 2017. We acquired our pedestrian and bicycle crash dataset from Minnesota Department of Public Safety (Department of Public Safety, 2018). The DPS dataset is compiled from police reports and includes all crashes for which a police report was filed. Some crashes, including some crashes that involve injury and visits to hospitals, are not reported to police. Hence, our dataset and analyses do not include all pedestrian or bicycle crashes that have occurred in Minneapolis during the period of interest. We analyze only pedestrian and bicycle crashes, which we define simply as a crash that involved pedestrian or bicyclist as coded in the DPS dataset. Given definitions in documentation, the dataset theoretically includes crashes between pedestrian and bicyclists that did not involve vehicles. However, none of these exist for these years, and all the crashes analyzed here involve vehicles. The dataset includes crash location, level of severity, time of the crash, and other details regarding the circumstances of the crash. We analyzed all pedestrian and bicycles crashes that occurred between 2005 and 2017 (13 years). We chose a long period of analyses because the longer period provides more observations and better insight into the spatial distribution of crashes. A tradeoff associated with the use of a longer period of analyses is that some factors associated with crashes may change over time, thus confounding or limiting the validity of correlation analyses.

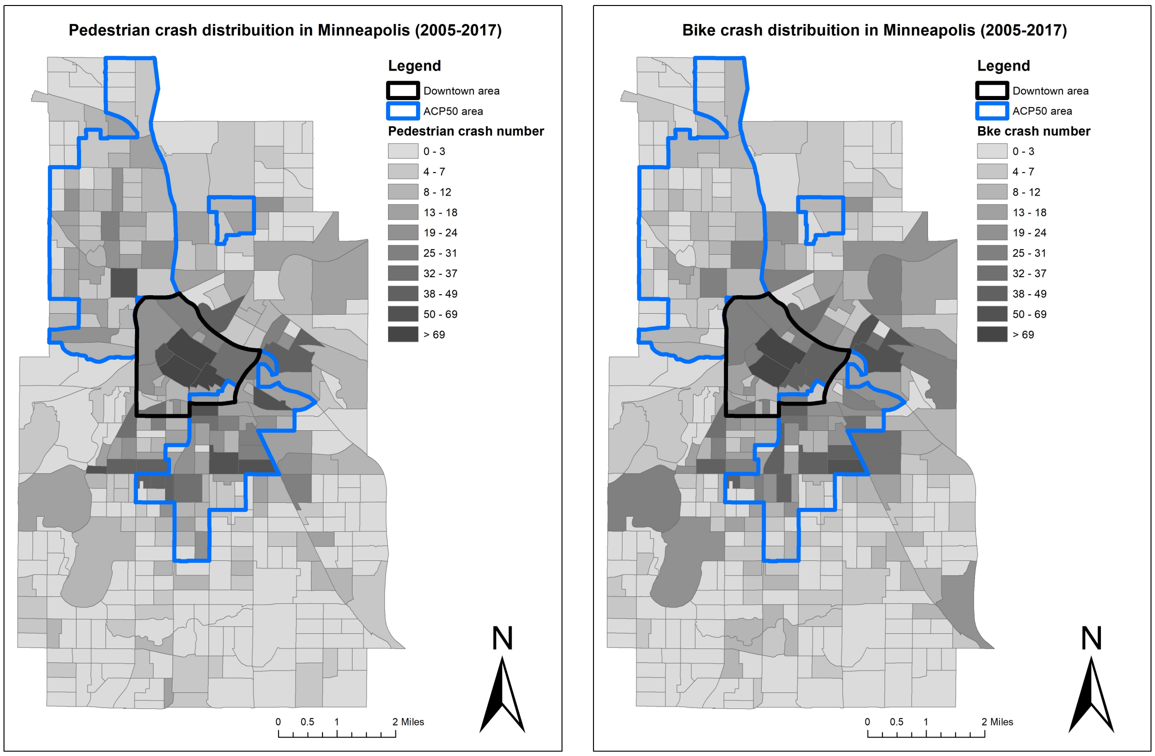
Table 3.1 and Figure 3.2, respectively, summarize the distribution of crashes and show the ranges in numbers of crashes in each census block group in three spatial areas in the city:

- The Minneapolis central business district, or downtown;
- Areas of Concentrated Poverty (ACP50) designated by the Metropolitan Council. These are census tracts in which 50% or more of residents are people of color, and 40% or more of the residents have family or individual incomes that are less than 185% of the federal poverty threshold. The Metropolitan Council and cities in the region, including Minneapolis, often consider these areas when spatially targeting equity initiatives.
- Other areas in Minneapolis (i.e., areas that are not in the CBD and designated as APC50).

Between 2005 and 2017, 3,812 pedestrian crashes and 3,490 bicycle crashes occurred in Minneapolis (Table 3.1). The crashes are not distributed evenly across areas within the city (Table 3.1). The density of crashes is highest in the CBD, and the density of crashes in the APC50 is higher than in other more affluent, majority-white areas of the city. Pedestrian crashes have slightly higher densities than bicycle crashes across in the CBD and APC50 tracts, but not in other areas of the city.

**Table 3.1 Pedestrian and bicycle crashes in different areas in Minneapolis**

| | City of Minneapolis | Central Business District Area | ACP50 Tracts | All Other Areas |
|---|---|---|---|---|
| Area (square miles) | 57.4 | 3 | 12.3 | 42.3 |
| Pedestrian crashes (2005-2017) | 3,812 | 904 | 1,490 | 1,443 |
| Pedestrian crash density (per sq mi) | 66.4 | 301.3 | 121.1 | 34.1 |
| Bike crashes (2005-2017) | 3,490 | 677 | 1,164 | 1,684 |
| Bike crash density (per sq mi) | 60.8 | 225.7 | 94.6 | 39.8 |

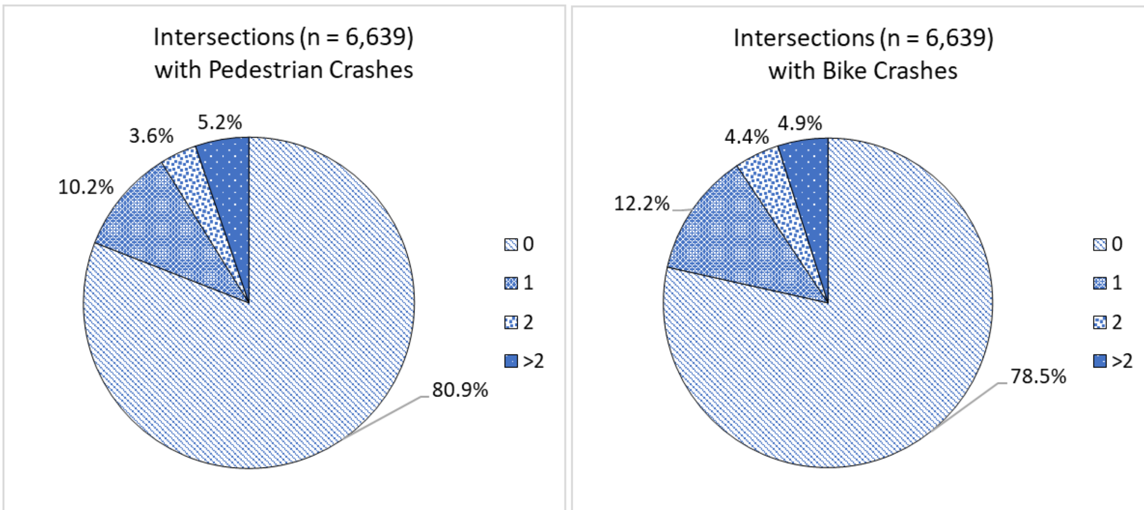



(a) (b)

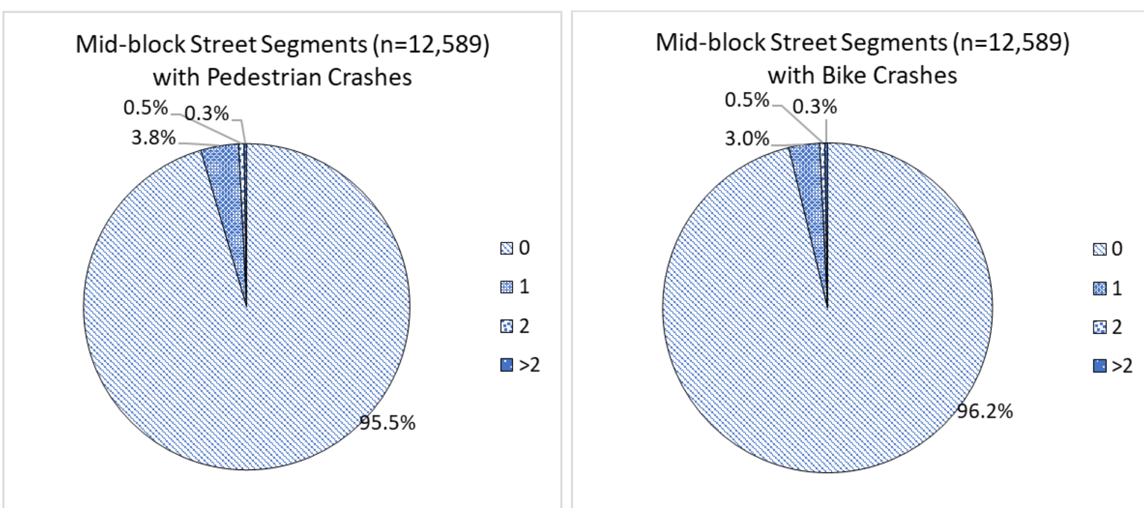
**Figure 3.2 Distribution of pedestrian (a) and bicycle (b) crashes (2005-2017) in each block group in Minneapolis. Note: the dataset for ACP50 area boundaries was extracted from the Minnesota Geospatial Commons (http://gisdata.mn.gov)**

We used ArcGIS to categorize each crash as either an intersection-related crash or a mid-block crash. Mid-blocks are those street segments that connect consecutive intersections. Our rationale for this distinction is that different factors may be associated with these two types of locations. We used different buffers for different types of roads to account for different road widths. We defined an intersection-related crash as a crash that occurred within a 35-meter buffer of the center of an intersection on major roads (e.g., arterials and collectors) and within a 15-meter buffer of the center of an intersection on local roads. We used the “intersect” function in ArcGIS to categorize or identify the intersection crashes. All crashes that were not classified as an intersection crash were classified as mid-block crashes; we used the “near” function in ArcGIS to assign these crash to the nearest mid-block.

We next identified the intersections and mid-blocks where one or more crashes occurred.  Figure 3.3 summarizes the numbers and percentages of intersections or mid-blocks where zero, one, or multiple crashes occurred during the period of analyses.  No pedestrian or bicycle crashes occurred at the vast majority of intersections (more than 78%) or mid-blocks (more than 95%).  Only about five percent of the intersections experienced three or more pedestrian or three or more bicycle crashes. Three or more crashes occurred at less than one percent of the mid-block locations.




(a)




(b)

**Figure 3.3 Number of intersections (a) and mid-blocks with zero, one, two, or three more crashes between 2005 and 2017**

We also categorized crashes based the severity of injury (Figure 3.4). Categories include fatality or killed, severe injury, moderate injury, minor injury, and property damage only. A majority of both pedestrian (51.6%) and bicycle (56.9%) crashes resulted in minor injuries. Fewer than 2% of both pedestrian (1.5%) and bicycle (0.5%) crashes resulted in fatalities.

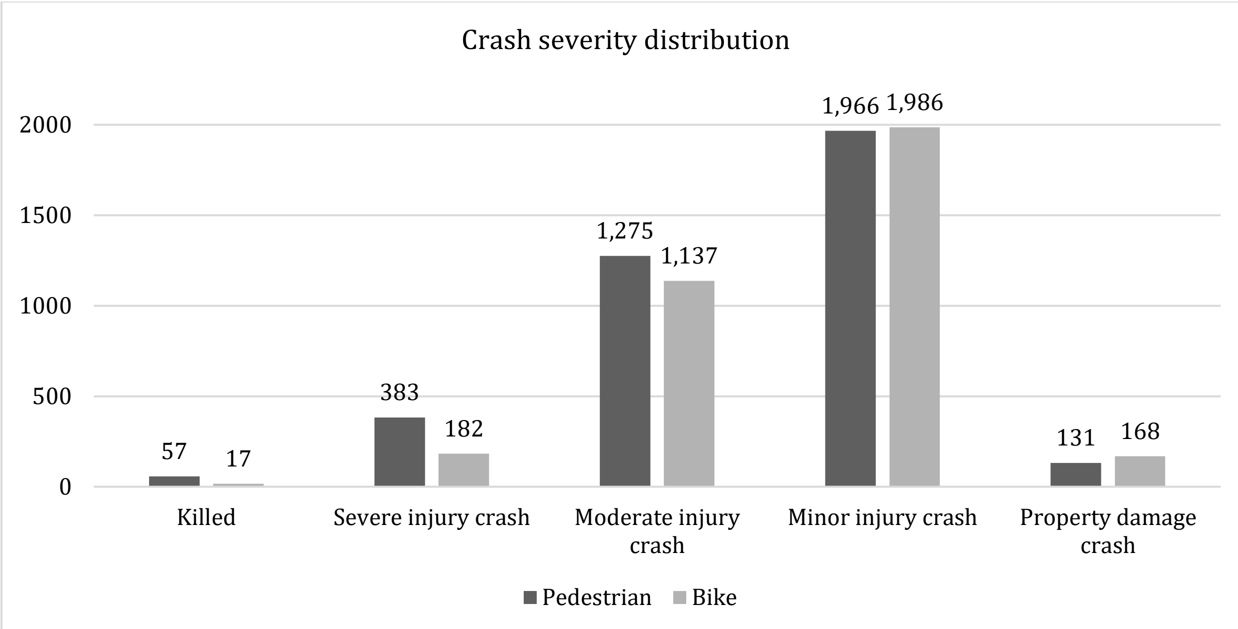



**Figure 3.4 Crash severity distribution**

## 3.2 MEASURES OF EXPOSURE TO RISK

As noted in our literature review, analyses of pedestrian and bicycle crashes historically have been hampered by the lack of measures of pedestrian and bicyclist exposure to risk. Although vehicle counting programs provide relatively fine-grained or disaggregate measures of vehicular exposure to risk, most communities have not implemented pedestrian or bicycle traffic monitoring programs to produce counts from which measures of exposure can be estimated. The Minneapolis DPW is distinctive among transportation agencies in major cities in the U.S. in that it has counted peak-hour pedestrian and bicycle traffic at multiple locations in the city for a number of years. A novel aspect of our research is that we use demand models estimated from DPW pedestrian and bicycle counts to produce estimates of pedestrian and bicyclist exposure to risk for every intersection and mid-block in the city (Hankey and Lindsey, 2016; Lindsey et al., 2018). These estimates, as will be explained below, are used with motorized traffic counts as measures of exposure to risk in our crash models

Minneapolis DPW Bicycle and Pedestrian Counts. The Minneapolis DPW, in collaboration with the nonprofit Transit for Livable Communities, initiated a manual, peak-hour pedestrian and bicycle monitoring program based on protocols developed by the National Bicycle and Pedestrian Documentation Project in 2007. These protocols call for counting bicyclists and pedestrian at mid-block locations in good weather in fall (mostly September) from 4:00 p.m. to 6:00 p.m. (Minneapolis Bicyclist and Pedestrian Count Report). For our measures of exposure, we use counts taken between 2007 and 2014, a time period that roughly corresponds to the time period for our crash data. We identified 437 unique mid-block locations where counts had been taken (Figure 3.5). For sites at which more than one count had been taken, we averaged counts to obtain a single value for analysis. These DWP counts are used directly as measures of exposure in our mid-block crash models.

Our measures of pedestrian and bicycle exposure have several limitations. One limitation is that two-hour, peak-hour weekday counts do not capture the variation in pedestrian and bicycle volumes that occurs throughout the day, by day-of-week, across seasons, or in response to weather. The time periods for the measures of exposure therefore do not match the time periods in the crash dataset (which includes crashes at all times of the day on all days of the week throughout the year). Another limitation is specific to analyses of mid-block pedestrian crashes. The DPW pedestrian counts measure pedestrians on sidewalks parallel to the street or roadway, not street crossings. The number of pedestrian crossings at each mid-block likely is substantially smaller than the number of pedestrians on the sidewalk. Our rationale for using these measures as estimates or proxies for exposure is that we believe them to be correlated with, and follow a similar distribution as, the desired measures (i.e., total pedestrian and bicycle volumes or mid-block crossing volume).

Because the DPW count database did not include intersection counts, we aggregated mid-block, or segment counts, to obtain intersection counts for analysis. We used the following procedures to aggregate segment counts to intersection counts for measures of exposure in our crash models. Specifically, we used Equation (1) to obtain measures of exposure for intersections linked with $N$ mid-blocks.

$$C_{intersection} = \frac{1}{2}\sum_{i=1}^{N} C_i \tag{1}$$

Where, $C_{intersection}$ is an intersection's pedestrian or bicycle count, and $C_i$ is the pedestrian and bicycle count of $i$th mid-block linked to the intersection.

We used Equation (2) to estimate counts for intersections if counts were not available for all mid-blocks leading to an intersection. For example, if one intersection had N mid-blocks linked to it, and M of them had actual counts, we estimated exposure as:

$$C_{intersection} = \frac{N}{2M}\sum_{i=1}^{M} C_i \tag{2}$$

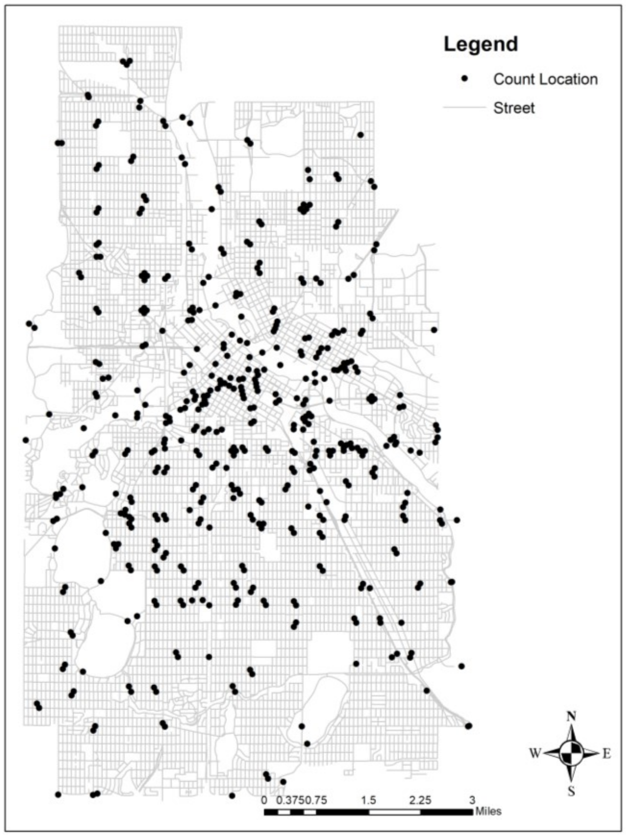



**Figure 3.5 Distribution of the count locations**

Use of these procedures resulted in counts, or measures of exposure, for 173 intersections. These counts were used to estimate crash models that subsequently were used to estimate crashes for all intersections. As a result of these procedures, our intersection and mid-block crash models have sample sizes of 173 and 437, respectively.

Because DPW identified the need to increase emphases on pedestrian and bicycle traffic and equity in its project ranking system, one of our research objectives was to estimate crash risk for every intersection (n=6,639) and mid-block (n=12,589) in the city. To do so required estimating pedestrian and bicyclist exposure to risk for all locations where actual counts were not available. To achieve this objective, we adapted estimates of pedestrian and bicycle peak-hour, mid-block traffic produced by pedestrian and bicycle demand models previously published by Hankey and Lindsey (2016). These models, which are based on counts at the 437 locations shown in Figure 3.5, estimate mid-block traffic as a function of adjacent land use, street functional class, and other variables (Hankey and Lindsey, 2016). After obtaining estimates for the 12,152 mid-blocks for which counts were not available by using these models, we then used equations 1 and 2 to obtain estimates for the remaining 6,466 intersections. A limitation of this method of estimating exposure at intersections is that we do not have actual measurements of crossings that enable detailed analyses of specific types of crashes.

For measures of vehicular exposure to risk, we obtained vehicle AADT of mid-blocks in Minneapolis from DPW. For mid-blocks or street segments where counts were not available, we used an estimated value of 500 passenger car units per day (pcu/day). We used Equation (1) to calculate the vehicle AADT for each intersection in Minneapolis.

## 3.3 CORRELATES OF CRASH RISK

Previous studies have shown that the probability or frequency of crashes is associated with characteristics of the roadway network, traffic controls, built environment, and socio-demographics of neighborhoods in addition to measures of exposure. Based on their findings, we assembled a number of different variables for the development of our crash models. The definitions and sources of the independent variables used in our models are listed in Table 3.2. Table 3.3 presents descriptive statistics (i.e., mean standard deviation) for these same variables for both intersections and mid-block locations.

Many characteristics of the built environment, roadway networks, and traffic control infrastructure are correlated, thus raising the potential for multi-collinearity in models. To address this potential, we constructed a correlation matrix of potential independent variables and excluded potential variables with correlation coefficients greater than 0.7 (e.g., % black population, which was highly correlated with % white population, and % other land use which was highly correlated with other land use categories. Figure 3.6 and Figure 3.7, respectively, show the correlations among the independent variables retained for use in our mid-block (n=437) and intersection (n=173) models. The darker shades represent higher correlations among variables; the lighter shades show weaker correlation. The matrixes also include the values of the correlation coefficients. None of the values exceeds 0.6, which is considered moderate correlation and acceptable for inclusion when modeling.

Inspection of our measures of exposure (i.e., estimates of vehicle, pedestrian, and bicycle volumes) revealed that distributions were skewed. Therefore, prior to modeling, we look the natural log of each measure to normalize its distributions. Figure 3.8 illustrates the distributions of pedestrian counts, bike counts, and AADT before and after taking the natural logarithm of each measure. As is evident in the graphs, the shape of each distribution after this procedure better approximates the normal distribution. For the measure of vehicular exposure (i.e., is AADT), one value has many more observations than others. This result occurs because, for residential and other low-volume streets where no counts have been taken, DPW assigns a standard AADT estimate of 500. Because no crashes have been reported on most of these mid-blocks, this limitation is not believed to have important effects on the modeling.

## 3.4 MODELS OF INTERSECTION AND MID-BLOCK CRASH RISK

We use regression analysis to model pedestrian and bicycle crash risk at mid-blocks (n=437) and intersections (n=173) where measures of exposure are available. Our dependent variables are the number of crashes at the intersections and mid-blocks, respectively. Because the distribution of crashes by location is positively (right-tail) skewed with many zeroes (i.e., no crashes have occurred at most mid-blocks and intersections), we use negative binomial regression models.

Our models include measures of exposure, the built environment, traffic control facilities, sociodemographic variables, and a binary, geographic variable for the CBD (Table 3.2 and Table 3.3). Each of our pedestrian and bicycle crash models includes three measures of exposure (i.e., pedestrian, bicycle, and motorized vehicle) because theory suggests interactions among all three modes of travel could affect the likelihood of crashes. Incorporation of these measures also is consistent with the DPW's goal of increasing emphases on pedestrian and bicycle traffic in its street project prioritization system. The other independent variables were selected based on previous research findings (Table 2.1).

We present two sets of pedestrian and bicycle crash models for intersections and mid-blocks. The first set comprises base models without sociodemographic variables that we use to assess the equity of distribution of crash risk between APC50 and other areas. Our rationale for excluding sociodemographic variables in our base models is that the spatial areas that we are comparing (i.e., the APC50 and other areas) were identified by the Metropolitan Council based on systematic differences in income and race. The inclusion of the socio-economic variables in the base models has the potential to confound results. Our second set of models, which we refer to as our final crash models, were estimated after completing this equity analysis. We added two variables, income and race, to our base models prior to estimating crashes at all intersections and mid-blocks throughout the city for project ranking. We do not analyze APC50 areas separately in this step in our analysis.

## 3.5 ASSESSING CRASH RISK AND EQUITY

We develop two measures to assess the equity of the spatial distribution of crash risk. The first is the average predicted number of crashes for different areas of the city, specifically, the CBD, the APC50 census tracts, and other areas. This measure is constructed separately for both pedestrian and bicycle crashes at both intersections and mid-blocks. The second measures, which also are computed separately for pedestrian and bicycle crashes at both intersections and mid-blocks, are Lorenz curves and GINI coefficients. These two indicators are measures of how far the distributions of crashes vary from perfect equality.

As noted in Section 3.4, we estimated base intersection (n=173) and mid-block (n=437) pedestrian and bicycle crash models without sociodemographic variables for purposes of testing for significant differences in indexes of crash risk between APC50 and other areas. After estimating the base models, we used them to predict crashes at all intersections and mid-blocks in the city. This procedure incorporated estimates of pedestrian, bicycle, and motorized traffic exposure as explained in Section 3.2. To assess equity of distribution of crash risk between the APC50, CBD, and other areas, we then averaged the crash numbers for all intersections and mid-blocks in each area, respectively, and conducted simple t-tests. As discussed in Chapter 4, this procedure revealed significant differences among areas, including significantly higher crash indexes in APC50 census tracts.

## Table 3.2 Variable definitions and their data sources

| Variables | | Definition | Data source | Year |
|---|---|---|---|---|
| | **Pedestrian or bicycle crash numbers** | Number of bicycle or pedestrian crashes occurs at an intersection or mid-block | Minnesota Department of Public Safety (DPS) | 2005-2017 |
| **Exposure** | **Actual pedestrian or bicycle counts** | Natural logarithm of PM peak hour (4-6 pm) bicycle or pedestrian counts | Minneapolis (DPW) | 2007-2014 |
| | **AADT** | Natural logarithm of vehicle average annual daily traffic (AADT) | DPW | 2005-2015 |
| **Built environment** | **Population density** | The number of people per acre of the block where an intersection or mid-block centroid is located | U.S. Census Bureau | 2010 |
| | **Job density** | The number of jobs per acre of the block where an intersection or mid-block centroid is located | U.S. Census Bureau | 2015 |
| | **Intersection number** | The number of intersections within 400-meter buffer of an intersection or mid-block centroid | DPW | - |
| | **Transit stop** | A dummy variable indicating whether there is a transit stop within 30-meter buffer of an intersection or mid-block. | Minnesota Geospatial Commons (MGC) | 2018 |
| | **Commercial area** | Percentage of commercial land use within 100-meter of an intersection or mid-block centroid. | MGC | 2016 |
| | **Office area** | Percentage of official land use within 100-meter of an intersection or mid-block centroid. | MGC | 2016 |
| | **Industrial area** | Percentage of industrial land use within 100-meter of an intersection or mid-block centroid. | MGC | 2016 |
| | **Open space** | Percentage of open space within 100-meter of an intersection or mid-block centroid. | MGC | 2016 |
| | **Land use entropy** | Land use entropy index within 100-meter buffer of an intersection or mid-block centroid | MGC | 2016 |
| **Geographical location** | **Downtown area** | A dummy variable indicating whether an intersection or mid-block is in the downtown area | DPW | 2018 |
| **Traffic facilities** | **Sidewalk** | A dummy variable indicating whether there is a sidewalk linked to an intersection or in a mid-block | DPW | - |
| | **Bicycle lane** | A dummy variable indicating whether there is a bicycle lane linked to an intersection or in a mid-block | DPW | 2011 |
| | **Trail** | A dummy variable indicating whether there is a trail linked to an intersection. | City of Minneapolis | 2014 |
| | **Street light** | A dummy variable indicating whether there is a street light in the 10-meter buffer of an intersection | DPW | 2018 |
| | **Traffic signal** | A dummy variable indicating whether there is a traffic signal in the 10-meter buffer of an intersection | DPW | 2018 |
| | **Main road** | A dummy variable indicating whether the mid-block is main road (e.g., arterials or collectors or other high volume roadways), or whether the intersection is linked with one main road | DPW | - |
| | **Secondary road** | A dummy variable indicating whether the mid-block is secondary road (e.g., a local, neighborhood, or other low-volume road), or whether the intersection is linked with one main road | DPW | - |
| **Socioeconomic** | **Child population** | Percentage of children (age 14 -) in total population of the block group where an intersection or mid-block centroid is located | U.S. Census Bureau | 2016 |
| | **Old population** | Percentage of old people (age 65 +) in total population of the block group where an intersection or mid-block centroid is located | U.S. Census Bureau | 2016 |
| | **Male population** | Percentage of male people in total population of the block group where an intersection or mid-block centroid is located | U.S. Census Bureau | 2016 |
| | **Average household size** | Average household size in square feet of the block group where an intersection or mid-block centroid is located | U.S. Census Bureau | 2016 |
| | **Average vehicle numbers** | Average vehicle numbers per household of the block group where an intersection or mid-block centroid is located | U.S. Census Bureau | 2016 |
| | **White population** | Percentage of white population in the block group where an intersection or mid-block centroid is located | U.S. Census Bureau | 2016 |
| | **Poverty population** | Percentage of poverty population in the block group where an intersection or mid-block centroid is located | U.S. Census Bureau | 2016 |

### Table 3.3 Descriptive statistics

| Variables | | Intersection (size = 173) | | Mid-block (size = 437) | |
|---|---|---|---|---|---|
| | | **Mean** | **Std. Dev.** | **Mean** | **Std. Dev.** |
| **Crash risk** | Pedestrian crash number | 3.13 | 5.16 | 0.24 | 0.83 |
| | Bicycle crash number | 2.21 | 3.03 | 0.24 | 0.75 |
| **Exposure** | Ln Pedestrian count | 6.72 | 1.20 | 5.91 | 1.27 |
| | Ln Bicycle count | 6.10 | 0.90 | 5.41 | 1.11 |
| | Ln AADT | 9.49 | 0.63 | 8.71 | 0.79 |
| **Built environment** | Population density | 14.16 | 24.22 | 13.71 | 19.81 |
| | Job density | 67.02 | 380.66 | 62.44 | 380.13 |
| | Intersection number | 29.13 | 7.83 | 28.92 | 8.50 |
| | Transit stop | 0.60 | 0.49 | 0.12 | 0.32 |
| | Commercial area | 0.21 | 0.25 | 0.17 | 0.24 |
| | Office area | 0.14 | 0.18 | 0.12 | 0.19 |
| | Industrial area | 0.04 | 0.13 | 0.05 | 0.17 |
| | Open space | 0.06 | 0.15 | 0.09 | 0.20 |
| | Land use entropy | 0.23 | 0.11 | 0.21 | 0.11 |
| **Downtown** | Downtown area | 0.17 | 0.37 | 0.16 | 0.37 |
| **Traffic facilities** | Sidewalk | 0.97 | 0.18 | 0.81 | 0.40 |
| | Bicycle lane | 0.64 | 0.48 | 0.21 | 0.41 |
| | Trail | 0.06 | 0.23 | - | - |
| | Street light | 0.77 | 0.42 | - | - |
| | Traffic signal | 0.66 | 0.47 | - | - |
| | Main road | 0.04 | 0.20 | 0.04 | 0.19 |
| | Secondary road | 0.61 | 0.49 | 0.41 | 0.49 |
| **Socio-demographic** | Child population | 0.15 | 0.10 | 0.15 | 0.11 |
| | Old population | 0.09 | 0.07 | 0.09 | 0.07 |
| | Male population | 0.52 | 0.08 | 0.52 | 0.07 |
| | Average household size | 2.23 | 0.67 | 2.26 | 0.69 |
| | Average vehicle number | 1.26 | 0.42 | 1.28 | 0.42 |
| | White population | 0.60 | 0.27 | 0.62 | 0.26 |
| | Poverty population | 3.24E-04 | 2.46E-04 | 3.22E-04 | 2.40E-04 |

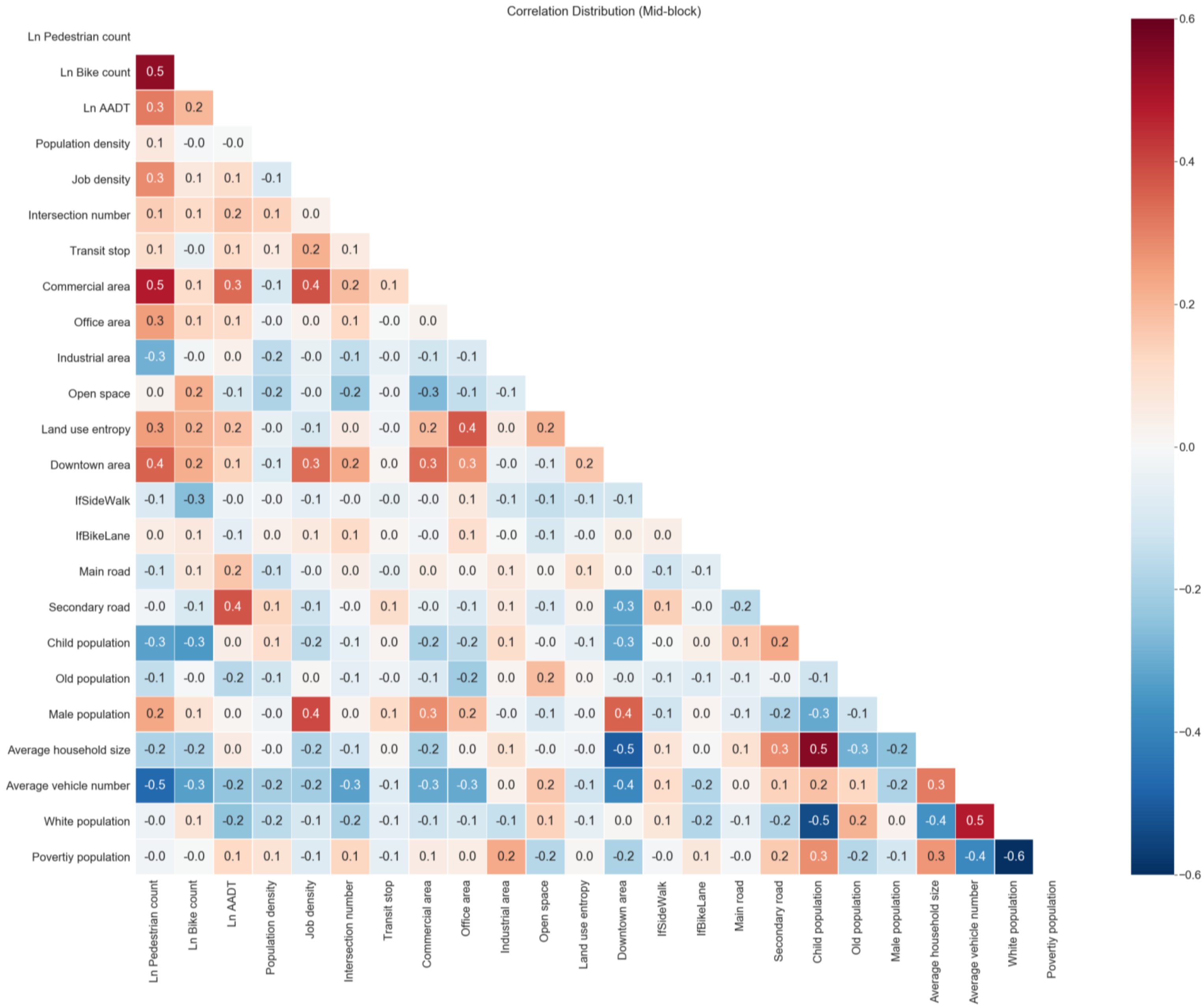



**Figure 3.6 Correlation distribution among variables of mid-blocks (n = 437)**

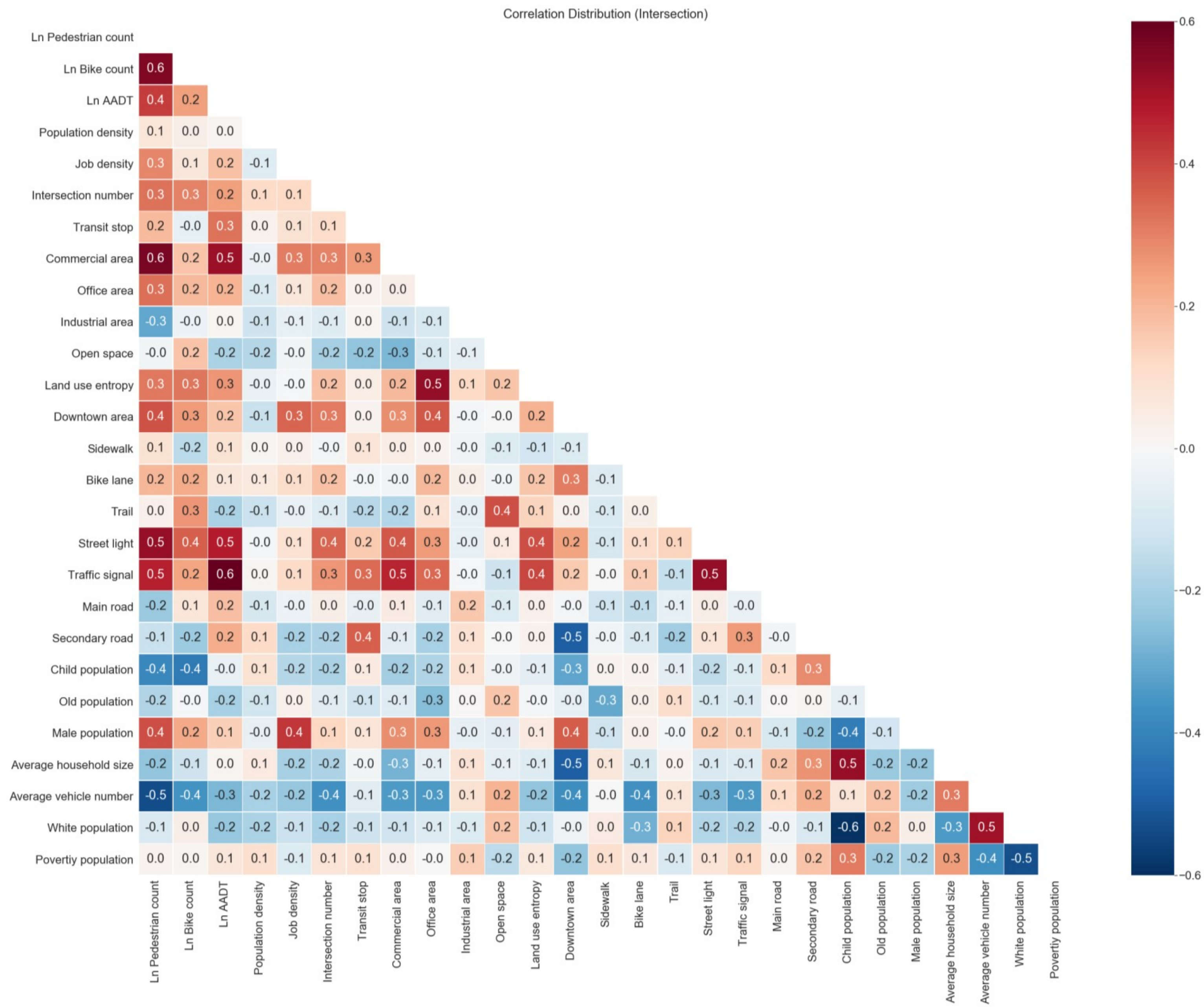



**Figure 3.7 Correlation distribution among variables of intersections (n = 173)**

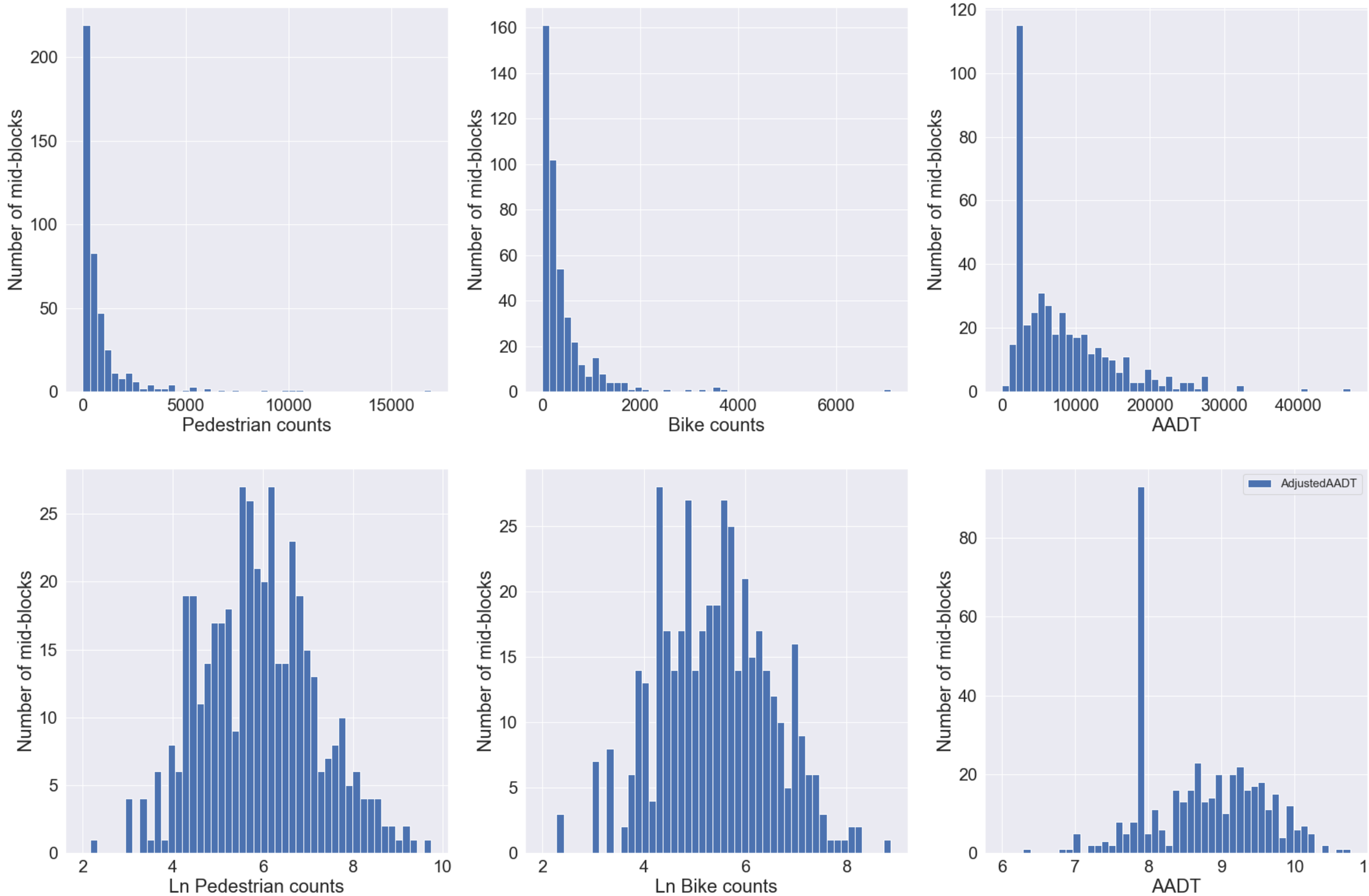



**Figure 3.8 Comparison between the distribution of bike, pedestrian counts, and AADT before and after natural logarithm**

Following this procedure, we re-estimated the base models, this time including independent variables for income and race, to obtain our final crash models. These final models then were used to predict pedestrian and bicycle crashes at all intersections and mid-blocks in the city (i.e., to produce our final crash indexes). Our rationale for incorporating the income and race variables in the final models is that, after controlling for other factors, the base models indicated that poverty and race likely are correlated with crashes.

We then used the crash indexes to construct Lorenz curves and Gini coefficients for both pedestrian and bicycle crashes at both intersections and mid-blocks for the entire city. Lorenz curves, which have been widely used in the field of economics, graphically represent the cumulative distribution of a resource across a population. The GINI coefficient is derived from Lorenz curves and take on a value from 0 to 1. The higher the value, the greater the inequality of the distribution. The GINI coefficient often is used to compare income inequality. For example, the GINI coefficient of income inequality in Minnesota is 0.45 while that in California is 0.49, indicating that the income distribution is somewhat more equal in Minnesota. Though their use in economics is most common, Gini coefficients also have been used in transportation to assess inequities in accessibility to bikeways (Wang and Lindsey, 2017). In this study, we used Lorenz curves and GINI coefficients to assess the distributions of pedestrian and bicycle crash risk. The Lorenz curve represents the relationship between the cumulative proportion of population and the cumulative proportion of crash risk to that population. We used the census block group as the areal unit of analysis for this procedure. We also calculated the corresponding GINI coefficients. Higher values of the GINI coefficient show greater inequality in the distribution of predicted crashes.

## 3.6 USING THE CRASH INDEXES IN PROJECT RANKING

The Minneapolis DPW uses an index, estimated crash rate, as a safety-related factor in its overall project ranking system. The DPW measure is constructed as the total number of all crashes at a location over a three-year period divided by the total estimated vehicular, transit, pedestrian, and bicycle traffic. Because vehicle-only crashes and vehicular traffic volumes are much higher than pedestrian and bicycle crashes and traffic volumes, this measure is weighted towards motorized traffic volumes, with the result that, all else equal, locations with the highest rate of vehicular crashes are prioritized. DPW does not apply its estimated crash rates directly in ranking. Instead, the crash rates are divided into four categories, and locations in each category are assigned points. Specifically, the factor for safety (i.e., the crash rate), is assigned up to 12 points in the overall ranking system. The points for the street average crash rate are awarded as follows:

- >5 crashes per million users per year: 12 points
- 2.5-4.9 crashes per million users per year: 8 points
- 1.0-2.5 crashes per million users per year: 4 points
- 0-0.9 crashes per million users per year: 0 points

This approach is an effective way to address safety-related concerns in project prioritization.

To address the DPW objective of increasing emphases on pedestrian and bicycle traffic, we explore the use of our pedestrian and bicycle crash indexes as a potential complement to the DPW indices. Specifically, we use our predicted crashes (for both pedestrians and bicycles at both intersections and mid-blocks), as an index, rank all locations from high to low, divide the ranked locations into quartiles, and assign points to locations in each quartile for purses of ranking. We then compare the ranking of locations (and potential projects) to the ranking based on the DPW crash rate and discuss the implications of the differences associated with these distinct, but related measures.

# CHAPTER 4: RESULTS AND DISCUSSION

We present our results in this chapter. We begin with the presentation of our base crash models (Section 4.1) that we use to predict crashes and assess the equity of the distribution of crash risk between APC50 and other areas in Minneapolis (Section 4.2). We then present our crash models for project ranking that include neighborhood income and race (Section 4.3). We apply these models to predict crashes at all intersections and mid-blocks in the city. These predicted values form our estimates of crash risk that we ultimately use to explore implications for project ranking. We present Lorenz curves and Gini coefficients to illustrate the extent to which the distributions of pedestrian and bicycle crashes at intersections and mid-blocks across the city depart from perfect equality (Section 4.4). Next, we illustrate how our crash indexes can be divided into quartiles, assigned points, and used to complement the crash rates used by DPW to rank or prioritize projects (Section 4.5). Specifically, we show how the use of new measures might change rankings based solely on the DPW estimated crash rate.

## 4.1 BASE CRASH MODELS

We estimated four base crash models: pedestrian intersection and mid-block models and bicycle intersection and mid-block models (Table 4.1). Each of the models includes exposure, built environment, traffic facility, the CBD, and socio-demographic variables. With the exception of three traffic facility variables, we include the same variables in each model to illustrate how pedestrian and bicycle crashes at intersections and mid-blocks are correlated with different variables. The pseudo $R^2$ values for the four models range from 0.13 to 0.31. The pedestrian intersection model has the best fit. These values cannot be interpreted as the percentage of variation in the dependent variable explained by the independent variables as can $R^2$ values used as goodness-of-fit statistics in ordinary least squares regression. The larger the value, the better the fit. The pseudo $R^2$ statistics show that the intersection models have better fit than the mid-block models. That is, more of the observed variation in the number of crashes is correlated with variation in the independent variables in the models. This result is consistent with prior expectations because mid-block crashes are rarer events.

A distinctive feature of these models is that we use vehicular (i.e., AADT), pedestrian, and bicycle counts to control for exposure in each model. This specification of the models yields interesting results. Our measure of vehicular exposure is significant in three of the models and nearly significant in the fourth model (i.e., the bicycle mid-block model; p=0.061; Table 4.1). The correlation in each model is positive, indicating that increased numbers of crashes are associated with higher vehicular traffic volumes. The significance of the pedestrian and bicycle exposure measures varies across the four models.

In the pedestrian intersection crash model, the pedestrian and vehicular measures of exposure both are correlated positively and significantly with the number of crashes. The bicycle exposure measure is inversely (negatively) correlated with the number of pedestrian crashes (Table 4.1). In the bicycle intersection crash model, both the bicycle and vehicular measures of exposure are positively correlated with the number of bicycle crashes, but the pedestrian exposure variable is not significantly correlated with the number of bike crashes. One hypothesis that could account for the inverse correlation between pedestrian crashes and bicycle exposure is that increased numbers of bicycles at intersections increases congestion, which leads to slower traffic and fewer conflicts and crashes between pedestrians and vehicles. The lack of significance of pedestrian exposure in the bicycle crash intersection model may be because some crashes between bicycles and vehicles occur at places in intersections outside of pedestrian crosswalks.

**Table 4.1 Base models for pedestrian and bicycle crashes at intersections and mid-blocks.**

| Variables | | Pedestrian models | | Bike models | |
|---|---|---|---|---|---|
| | | Intersection (size = 173) | Mid-block (size = 437) | Intersection (size = 173) | Mid-block (size = 437) |
| Exposure | Ln Pedestrian count | 0.67*** | 0.49* | 0.10 | 0.12 |
| | Ln Bike count | -0.27** | -0.05 | 0.54*** | 0.48** |
| | Ln AADT | 1.17*** | 0.84** | 0.52** | 0.43 |
| Built environment | Population density | 4.91E-03** | 0.02** | 4.52E-03 | 0.01 |
| | Job density | -3.01E-04 | 2.50E-04 | -9.26E-05 | -5.50E-04 |
| | Intersection number | 0.01 | 0.01 | -0.02 | 2.09E-03 |
| | Transit stop | 0.54** | -0.92 | -0.02 | -0.27 |
| | Commercial area | -0.47 | -0.27 | -0.11 | 1.72* |
| | Office area | -0.60 | -0.44 | -0.07 | -0.15 |
| | Industrial area | -1.14 | 0.71 | 0.56 | 1.54 |
| | Open space | -2.90*** | -0.27 | -1.29 | 2.22** |
| | Land use entropy | -0.29 | -0.55 | 1.04 | -0.22 |
| Geographical location | Downtown area | -0.24 | 0.30 | -0.34 | 1.17** |
| Traffic facilities | Sidewalk or bike lane | -0.48 | 0.80 | -0.06 | -0.16 |
| | Trail | 0.61 | - | -0.51 | - |
| | Street light | 0.14 | - | 0.34 | - |
| | Traffic signal | 0.68** | - | 0.43 | - |
| | Main road | 0.99*** | 0.66 | -0.58 | -1.14 |
| | Secondary road | 0.39 | 0.36 | 0.47 | 0.45 |
| Socioeconomic | Child population | 2.45** | 0.08 | 0.33 | -0.09 |
| | Old population | -4.39*** | 0.33 | -1.42 | 3.21 |
| | Male population | -0.90 | 0.57 | -0.75 | -0.18 |
| | Average household size | -0.48*** | -0.21 | -0.24 | 0.37 |
| | Average vehicle numbers | 0.01 | -0.39 | -0.14 | -0.13 |
| | Constant | -12.90*** | -12.53*** | -7.74*** | -10.78*** |
| Model fitness | Pseudo R2 | 0.31 | 0.13 | 0.18 | 0.13 |
| | Log likelihood | -261 | -213 | -282 | -225 |

Note: *P>z at 95% level; ** P>z at 99% level; *** P>z at 99.9% level

In the pedestrian mid-block model, in addition to vehicular exposure, pedestrian exposure is significantly correlated with pedestrian crashes. In the bicycle mid-block model, only bicycle exposure is significantly correlated with the number of bicycle crashes.  The lack of significance of the pedestrian exposure variable in the bike model (and vice-versa) may be because these modes are unlikely to interact or complicate interactions of other modes at mid-blocks.

We follow Elvik (2013) to interpret the implications of these results for the safety-in-numbers hypothesis. For the pedestrian intersection model (Table 4.1), the coefficient of ln-pedestrian count is positive but less than 1, which is consistent with the "safety in numbers" hypothesis and findings in previous papers (Elvik, 2013, 2009; Elvik and Bjørnskau, 2017; Schneider et al., 2010). At the same time, the coefficient on vehicular exposure (i.e., ln-AADT) is greater than one, indicating the presence of a "hazard-in-numbers" effect. Elvik (2013) characterizes this type of outcome (i.e., with both safety- and hazard-in-numbers in the same dataset) as "partial" as opposed to "complete" safety-in-numbers. All the coefficients in the three other models are less than one, which may be interpreted of existence of at least partial safety-in-numbers in the models. Elvik (2013, p. 58) cautions: "If pedestrian or cyclist volume is highly correlated with motor vehicle volume, there will be no overall safety-in-numbers effect with respect to total traffic volume if the sum of the coefficients is greater than 1." None of correlation coefficients among measures of exposure exceeded 0.6, which reflects only moderate correlation Figure 3.6 and Figure 3.7). Thus, the other three models potentially constitute evidence of complete safety-in-numbers effects. This evidence appears strongest in the bicycle mid-block model: the coefficients of the pedestrian, bicycle, and vehicular exposure measures sum to 1.03 (Table 4.1). The sums of coefficients for the exposure variables for the bicycle intersection and pedestrian midblock models, respectively, are 1.16 and 1.28 (Table 4.1).

We include nine different built environment variables in each base model (Table 4.1). The pedestrian model has the best fit overall. Comparatively few of the built environment variables are significantly associated with crashes overall. Pedestrian crashes at intersections are positively and significantly associated with population density and transit stops. Pedestrian intersection crashes are negatively and significantly associated with the percentage of nearby land in open space. Pedestrian mid-block crashes are positively and significantly correlated with population density. Bicycle crashes at intersections are not significantly correlated with any of the variables in the model other than bicycle and vehicle exposure. Mid-block bicycle crashes, however, appear to be associated with features of the built environment. These crashes are positively and significantly correlated with the percentage of nearby commercial area and the percentage of nearby land in open space.

Most of the significant correlations between crashes and elements of the built environment are in the hypothesized direction. For example, it is expected that more pedestrian intersection crashes will occur in areas with higher population density. More bicycle crashes may occur at mid-blocks near commercial areas because of interactions between drivers who are parking and cyclists (e.g., crashes associated with drivers opening doors as bicyclists are passing by). The inverse relationship between pedestrian intersection crashes and open space may be because pedestrians are more visible to drivers. Additional study is needed to explain the positive correlation between mid-block bicycle crashes and nearby open space.

A surprising outcome given the descriptive results presented in Section 3.1 that show the density of crashes is higher in the CBD is that the binary CBD variable is significant in only the bicycle mid-block model. This result may be because, after controlling for exposure and other variables, the underlying factors that affect the likelihood of crashes are not substantially different.

Only two of the traffic facility variables are significant in the four models. Pedestrian intersection crashes are significantly and positively associated with the main road variable, which includes arterials and collectors, and with the presence of traffic signals.

Pedestrian intersection crashes are more likely to be associated with nearby socio-demographic characteristics than other types of crashes. The percentage of children in nearby areas is positively and significantly associated with numbers of crashes while the percentage of elderly in nearby populations is negatively and significantly associated with pedestrian intersection crashes. Average household size also is inversely associated with both pedestrian intersection crashes.

## 4.2 PREDICTED CRASHES IN APC50 TRACTS, THE CBD, AND OTHER AREAS

We used our four base models to predict numbers of bicycle and pedestrian crashes for all intersections and mid-blocks in Minneapolis. We then averaged projected number of crashes for all intersections and mid-blocks, respectively in three areas with the city: Metropolitan Council designated APC50s, non-APC50 areas, and the CBD. We used simple t-tests to determine if the predicted crashes for the different areas are significantly different.

All neighborhoods in the city (i.e., census block groups) have been classified as either APC50 or non-APC50. The CBD, however, includes both APC50 and non-APC50 areas. As shown in Figure 3.2, the density of crashes in the CBD is much higher. We therefore first test APC50 versus non-APC50 areas. Then, we remove the subareas that are within the CBD and repeat the test. The rationale for this approach is that APC50 and non-APC50 areas within the CBD are more to be similar and, potentially, mask differences between the areas.

Overall, the results show that, on average, predicted pedestrian and bicycle intersection crashes are significantly higher in APC50 than in non-APC50 areas. With respect to mid-block crashes, differences depend on whether the CBD is included or excluded from the test. When the CBD is included, there are no significant differences in crash risk between APC50 and non-APC50 areas. However, when the CBD is excluded, both pedestrian and bicycle mid-block crash risk is significantly higher in APC50 areas

To illustrate the magnitude of differences, we can compare differences in crash indexes. The mean APC50 pedestrian and bicycle intersection crash indexes are 31% and 25% higher, respectively, than the comparable measures for the non-APC50 areas (Table 4.2). The differences in mean pedestrian and bicycle crash indexes for mid-blocks between APC50 and non-APC50 areas are, respectively, 0% and 6% (Table 4.3).

The differences in mean predicted crashes between APC50 and non-APC50 areas become even more pronounced when the CBD is not included in the comparisons. The mean APC50 pedestrian and bicycle intersection crash indexes are 80% and 38% higher, respectively, than the comparable measures for the non-APC50 areas (Table 4.4). The differences in mean pedestrian and bicycle crash indexes for mid-blocks between APC50 and non-APC50 areas are, respectively, 35% and 25% (Table 4.5).

People in Minneapolis who live in lower-income neighborhoods in which more than half the population is minority face higher crash risk than those individuals who live in more affluent, majority-white neighborhoods, especially at intersections.

**Table 4.2 T-test result between bicycle and pedestrian crash number at intersections (including CBD)**

| Intersection | ACP50 area mean | Non-ACP50 area mean | Difference | P-value |
|---|---|---|---|---|
| Pedestrian | 0.4381 | 0.3349 | 31% ↑ | 0.0018 |
| Bicycle | 0.3164 | 0.2535 | 25% ↑ | 0.0001 |

**Table 4.3 T-test results between bicycle and pedestrian crash number in mid-blocks (including CBD)**

| Mid-block | ACP50 area mean | Non-ACP50 area mean | Difference | P-value |
|---|---|---|---|---|
| Pedestrian | 0.0484 | 0.0481 | 0% - | 0.9388 |
| Bicycle | 0.0419 | 0.0447 | 6% ↓ | 0.1823 |

**Table 4.4 T-test results between bicycle and pedestrian crash number at intersections (without CBD)**

| Intersection | ACP50 area mean | Non-ACP50 area mean | Difference | P-value |
|---|---|---|---|---|
| Pedestrian | 0.4404 | 0.2453 | 80% ↑ | 0 |
| Bicycle | 0.3159 | 0.2286 | 38% ↑ | 0 |

**Table 4.5 T-test results between bicycle and pedestrian crash number in mid-blocks (without CBD)**

| Mid-block | ACP50 area mean | Non-ACP50 area mean | Difference | P-value |
|---|---|---|---|---|
| Pedestrian | 0.0460 | 0.0341 | 35% ↑ | 0 |
| Bicycle | 0.0405 | 0.0324 | 25% ↑ | 0 |

## 4.3 CRASH MODELS FOR PROJECT RANKING

We present in Table 4.6 the crash models we use to develop indexes to inform ranking of street improvement projects. As noted in Sections 3.4 and 4.1, our base models did not include income and race variables because we estimated them specifically to compare APC50 and non-APC50 areas that are defined on the basis of income and race. These two variables are included in the models presented in Table 4.6 so that these factors are controlled for in the final estimation process.

**Table 4.6 Crash Models for Project Ranking (with racial and poverty variables)**

| Variables | | Pedestrian models | | Bike models | |
|---|---|---|---|---|---|
| | | Intersection (size = 173) | Mid-blocks (size = 437) | Intersection (size = 173) | Mid-blocks (size = 437) |
| **Exposure** | Ln Pedestrian count | 0.76*** | 0.52* | 0.12 | 0.11 |
| | Ln Bike count | -0.27** | -0.05 | 0.56*** | 0.48** |
| | Ln AADT | 1.15*** | 0.84** | 0.44* | 0.43 |
| **Built environment** | Population density | 0.01** | 0.02** | 3.83E-03 | 0.01 |
| | Job density | -3.48E-04* | 2.33E-04 | -1.32E-04 | -5.32E-04 |
| | Intersection number | 1.95E-03 | 0.01 | -0.02 | 2.19E-03 |
| | Transit stop | 0.52** | -0.86 | -0.05 | -0.28 |
| | Commercial area | -0.60* | -0.35 | -0.16 | 1.76* |
| | Office area | -0.46 | -0.45 | -0.06 | -0.14 |
| | Industrial area | -1.01 | 0.68 | 0.59 | 1.57 |
| | Open space | -3.02*** | -0.25 | -1.44* | 2.24** |
| | Land use entropy | -0.07 | -0.49 | 1.00 | -0.26 |
| **Downtown** | Downtown area | -0.36 | 0.32 | -0.45 | 1.18** |
| **Traffic facilities** | Sidewalk or bike lane | -0.39 | 0.79 | -0.10 | -0.16 |
| | Trail | 0.63 | - | -0.43 | - |
| | Street light | -0.05 | - | 0.27 | - |
| | Traffic signal | 0.80** | - | 0.50 | - |
| | Main road | 1.22*** | 0.77 | -0.46 | -1.16 |
| | Secondary road | 0.30 | 0.31 | 0.51* | 0.47 |
| **Socioeconomic** | Child population | 0.70 | -0.82 | -0.94 | 0.34 |
| | Old population | -4.77*** | 0.34 | -1.74 | 3.23 |
| | Male population | -1.33 | 0.44 | -1.41 | -0.08 |
| | Average household size | -0.71*** | -0.28 | -0.31 | 0.40 |
| | Average vehicle numbers | 0.69* | -0.07 | 0.03 | -0.26 |
| | White population | -1.32** | -0.49 | -1.00 | 0.34 |
| | Poverty population | 220.37 | 294.29 | -601.93 | 48.94 |
| | Constant | -12.23*** | -12.49*** | -5.91** | -11.04*** |
| **Model fitness** | Pseudo R2 | 0.32 | 0.13 | 0.19 | 0.13 |
| | Log likelihood | -256 | -213 | -280 | -225 |

Note: *P>z at 95% level; ** P>z at 99% level; *** P>z at 99.9% level

Overall, these models are quite similar to the base models, although some changes in the significance of variables occur. More variables become significant in the pedestrian intersection model: job density (negative), percentage of commercial area (negative), average number of vehicles (positive, and percentage white population (negative). The negative correlations between job density and commercial areas, respectively, and pedestrian crashes is the opposite of what is expected, but may because another variable (i.e., downtown area) controls for the CBD. Evidence of both safety-in-numbers and hazard-in-numbers is retained in the pedestrian intersection model. An interesting outcome is that the race variable (i.e., percent white population) is significant in only pedestrian intersection model and the income variable is not significant in any of the four models (Table 4.6). Pedestrian crashes at intersections are negatively and significantly correlated with the percentage of the nearby population that is white. The lack of significance of the race and income variables in the other models may be associated with the relatively small area (census block group) for which these variables were measured.

We used our crash models to predict crashes for every intersection (n= 6,639) and mid-block (n= 12,589) in Minneapolis. To illustrate the distribution of predicted crashes, we divided values into quartiles and prepared heat maps of these quartiles (Figure 4.1). These maps show that predicted crashes are highest in the CBD and along arterials throughout the city. A comparison of these maps with actual crash locations (Figure 3.1) shows that patterns generally are similar. One value of these maps is that they illustrate potential risk at locations where no crashes ever have occurred. This feature of our approach enables comparison and ranking of these sites in a systematic way.

We also constructed quantile-quantile (q-q) plots to compare the distribution of predicted crashes at intersections and mid-blocks to their corresponding, actual values (Figure 4.2). The dashed line represents perfect prediction, or correlation, between predicted and actual crashes. These plots show that the predicted and actual crashes have similar density distributions, but differences emerge at the upper ends of the distributions. For the intersection plots, the predicted values are higher than actual values at sites with the most crashes, while for the mid-block plots, the predicted values appear to be closer. Because our purpose in constructing these measures is to rank and prioritize locations, the absolute differences in magnitude between predicted and actual are not as important as they might be with other applications.

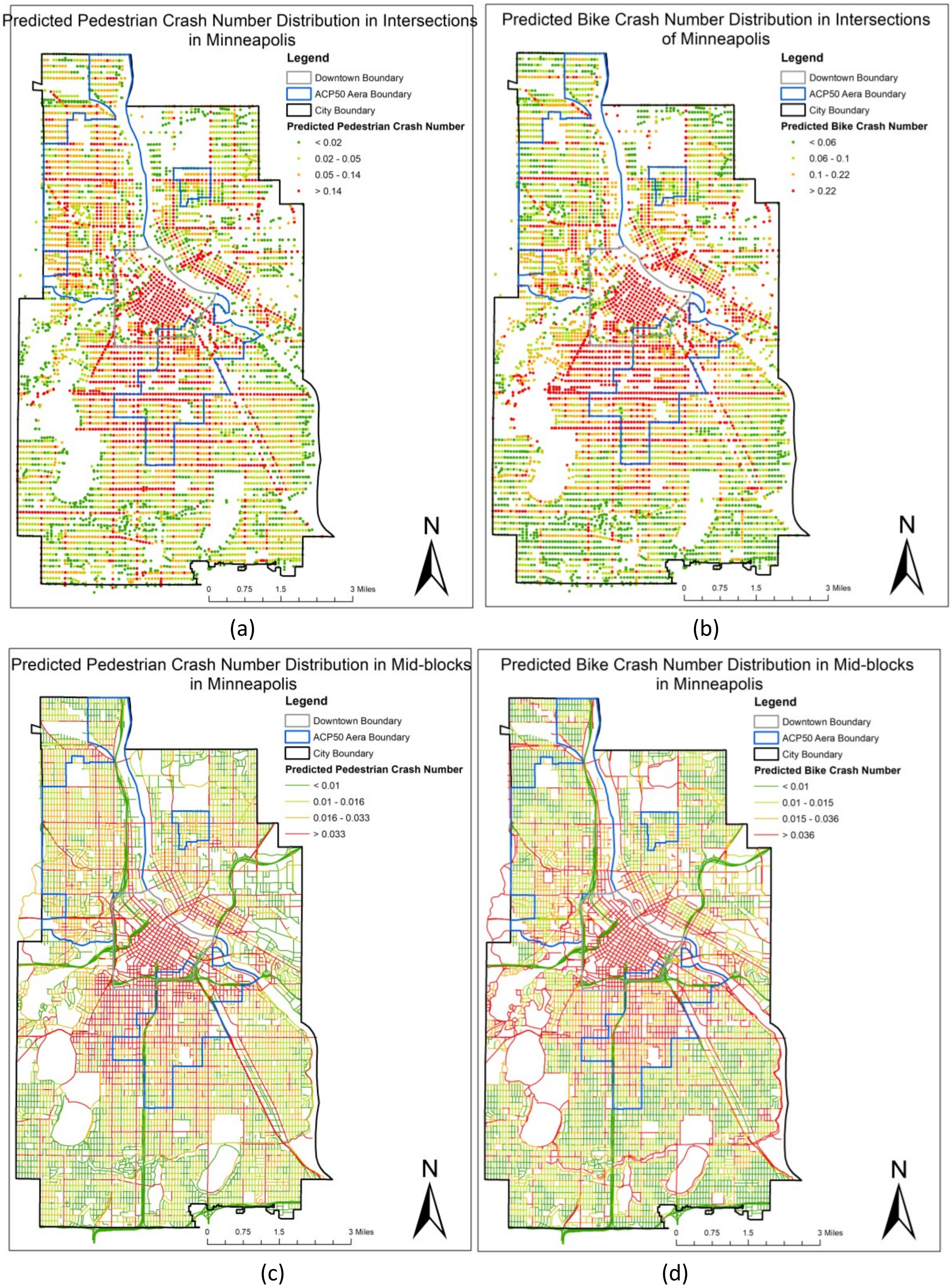



**Figure 4.1 Crash numbers distribution in Minneapolis: pedestrian and bicycle crash numbers at intersections and mid-blocks**

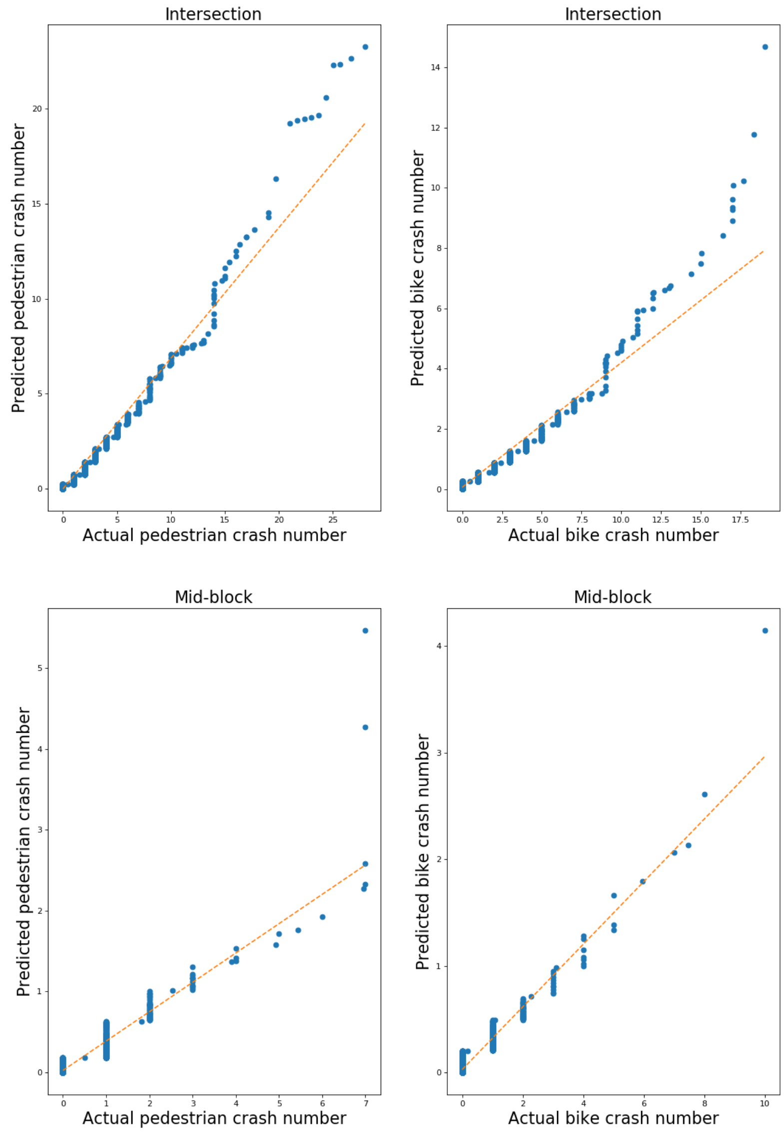
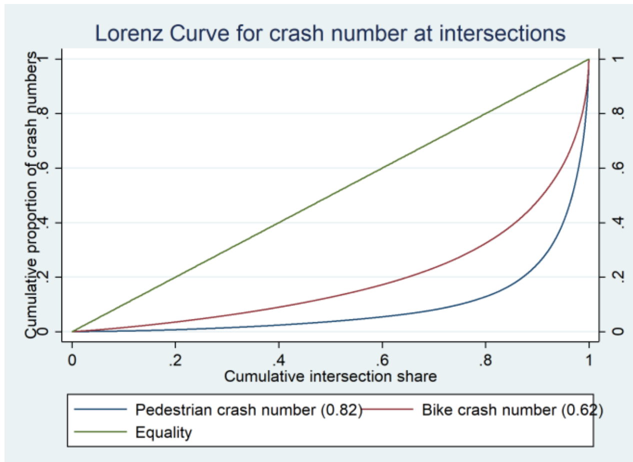


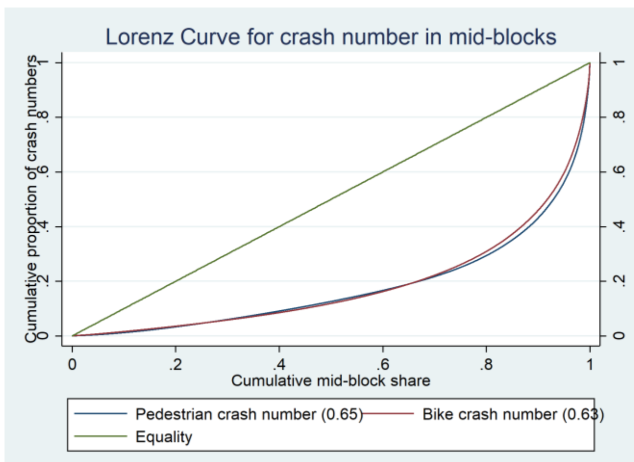

**Figure 4.2 Quantile-Quantile plots for actual crash number v. predicted crash number (the yellow line is linearly fitted line)**

## 4.4 EQUITY OF DSITRIBUTION OF PREDICTED CRASHES

The statistical tests in Section 4.2 and the heat maps in Figure 4.1 show clearly that crash risk, as measured by predicted crashes, is not distributed evenly throughout Minneapolis. To provide additional insight into the distribution of crash risk, we plotted Lorenz curves and calculated Gini Coefficients for pedestrian and bicycle intersection and mid-block predicted crashes, respectively (Figure 4.3). These plots show that (a) the distributions of intersection crash risk depart more from perfect equality than do the distributions of mid-block crash risk, and (b) predicted pedestrian crashes are less evenly distributed than predicted bicycle crashes. A useful feature of these plots is that they can be used to estimate relative shares, or burden, of risk. For example, 20% of intersections account for approximately 85% of predicted pedestrian intersection crashes. Predicted bicycle crashes are more evenly distributed: approximately 40% of intersections account for 85% of predicted crashes. Patterns are similar for mid-blocks predicted crashes, but differences between pedestrians and bicycle crashes are smaller.




(a)




(b)

**Figure 4.3 Lorenz curves for predicted pedestrian and bicycle crashes at intersections (a) and mid-blocks (b)**

## 4.5 CRASH RISK INDEXES FOR PROJECT RANKING

An important objective of this research is to develop measures of pedestrian and bicycle crash risk and equity that can be used to complement or augment criteria being used by the Minneapolis DPW to rank street improvement projects. As noted in Section 3.6, the DPW presently uses estimated crash rates as indexes to assign points and prioritize safety in its ranking system. Locations with higher crash rates are assigned higher point totals, indicating higher priority. Crash rates are grouped into four levels; locations with the highest crash rates received 12 points, and, in descending order, the three lower levels receive 8, 4, or 0 points. This approach is effective in addressing safety but it has limitations. One limitation is that numbers of vehicle-only crashes are substantially greater than numbers of pedestrian and bicycle crashes. As a result, the numbers of bicycle and pedestrian crashes do not greatly influence overall project rankings. A second limitation is that because the rate is calculated using actual crash numbers, and because no crashes have occurred at many sites, many locations have no crash risk. Although the DWP addresses this limitation in project selection by selecting locations with maximum values along corridors, risk at locations where no crashes have occurred is not considered directly in ranking. We illustrate here how our predicted crash numbers can be used as an index to assign points with the DPW ranking system.

Thus far, we have presented separate models and analyses of predicted pedestrian and bicycle intersection and mid-block crashes. To simplify incorporation of our results in a ranking example, we combine measures from the four models into a single average for purposes of assigning points. Specifically, we average our predicted crash values to create a combined pedestrian and bicycle crash index for purposes of assigning points. We also illustrate how we can adapt Gini coefficients to inform ranking.

### 4.5.1 Combined Pedestrian and Bicycle Crash Index

We calculated our combined pedestrian and bicycle crash index for all mid-blocks in the city. The reason for aggregating to mid-blocks is to be consistent with DPW procedures. Figure 4.4 depicts graphically how the predicted pedestrian and bicycle crashes at intersections and mid-blocks were averaged to produce this index. In typical, but not all cases, each mid-block is a street segment between two intersections. Each intersection and mid-block has two predicted crash numbers. In Figure 4.4, in total six predicted crash numbers are associated with the two intersections and single mid-block. In this example, we would average these six numbers and use this average as the combined pedestrian and bicycle crash index.

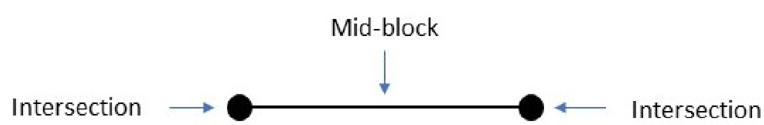



**Figure 4.4 Construction of combined pedestrian and bicycle crash index for one mid-block and two intersections**

To be consistent with current DPW procedures, we next divided the combined index into quartiles and assigned points based on those quartiles, using the same point totals as DPW uses in assigning points for safety (Table 4.7). To illustrate the uneven nature of the distribution of the combined index, we plotted the quartiles (Figure 4.5). The plot confirms that a relatively small proportion of all mid-blocks in the city have high index values (or, more practically, comparatively high risk). To assess the validity of our combined crash index, we test the correlation between its value and the corresponding average actual crash numbers (i.e., the total number of pedestrian and bicycle crashes that occurred in the intersections and mid-block included in the corresponding index. The correlation coefficient is 0.47, which indicates moderate correlation.

**Table 4.7 Points awarded for different groups with different averaged predicted crash number**

| Averaged predicted crash number | Median value | Points awarded |
|---|---|---|
| Quantile 1: 0.0 - 0.04 | 0.03 | 0 |
| Quantile 2: 0.04 - 0.06 | 0.05 | 4 |
| Quantile 3: 0.06 - 0.16 | 0.09 | 8 |
| Quantile 4: 0.16 – 9.05 | 0.39 | 12 |

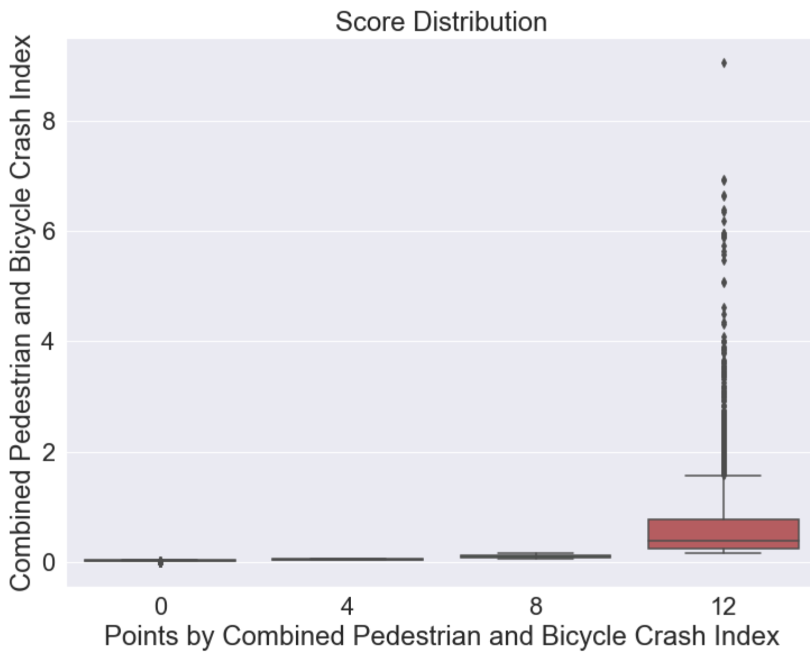



**Figure 4.5 Distribution of points given by averaged predicted crash number**

### 4.5.2 Gini Coefficient for Combined Crash Index

We used Gini Coefficients to demonstrate how pedestrian and bicycle crashes at intersections and mid-blocks are unevenly distributed throughout the city of Minneapolis (Section 4.4). Analyses of evenness in distribution of phenomena of interest – in this case crash indexes – vary depending on the areal unit of analysis. Within more fine-grained areal units, different patterns or distributions may emerge. To illustrate this phenomenon, and to provide an example of now unevenness in predicted crashes in smaller project areas could be used to rank projects, we calculated Gini coefficients for each census block group in the city using our combined crash index. This procedure involved determining the number of mid-blocks within each census block group and then using the combined crash indices to compute the Gini coefficients for each block group. The distribution of Gini coefficients across census block groups is shown in Figure 4.6.

Table 4.8 presents point totals associated with each quartile, again using overall point totals consistent with DPW values for safety. The idea here is that prioritization of projects in small areas with wide disparities in crash risk may be in the public interest.


Figure 4.6 Distribution of points given by crash risk Gini coefficient

Table 4.8 Points awarded for different groups with different crash risk Gini coefficient

| Averaged predicted crash number | Median value | Points awarded |
|---|---|---|
| Quantile 1: 0.0 - 0.32 | 0.24 | 0 |
| Quantile 2: 0.32 - 0.45 | 0.39 | 4 |
| Quantile 3: 0.45 - 0.56 | 0.50 | 8 |
| Quantile 4: 0.56 - 0.78 | 0.61 | 12 |

### 4.5.3 Use of the Combined Crash Indexes in Project Ranking

In this section, we present two examples of how our indices can be used to revise or augment the DPW ranking systems. The first example involves replacing the safety index (i.e., crash rate) used by DPW with the two new indices and comparing the two distributions. The second involves adding the two new indexes to the safety index and comparing the three-dimensional index to the current safety index. Because all measures are related to crashes, we computed correlation coefficients between the DPW safety index and the combined crash index and between the DPW safety index and our census block group Gini coefficient index. The two correlation coefficients were, respectively, 0.25 and 0.16, which indicate a weak correlation. The implication is that ranking systems based on these different measures will result in different project prioritization.

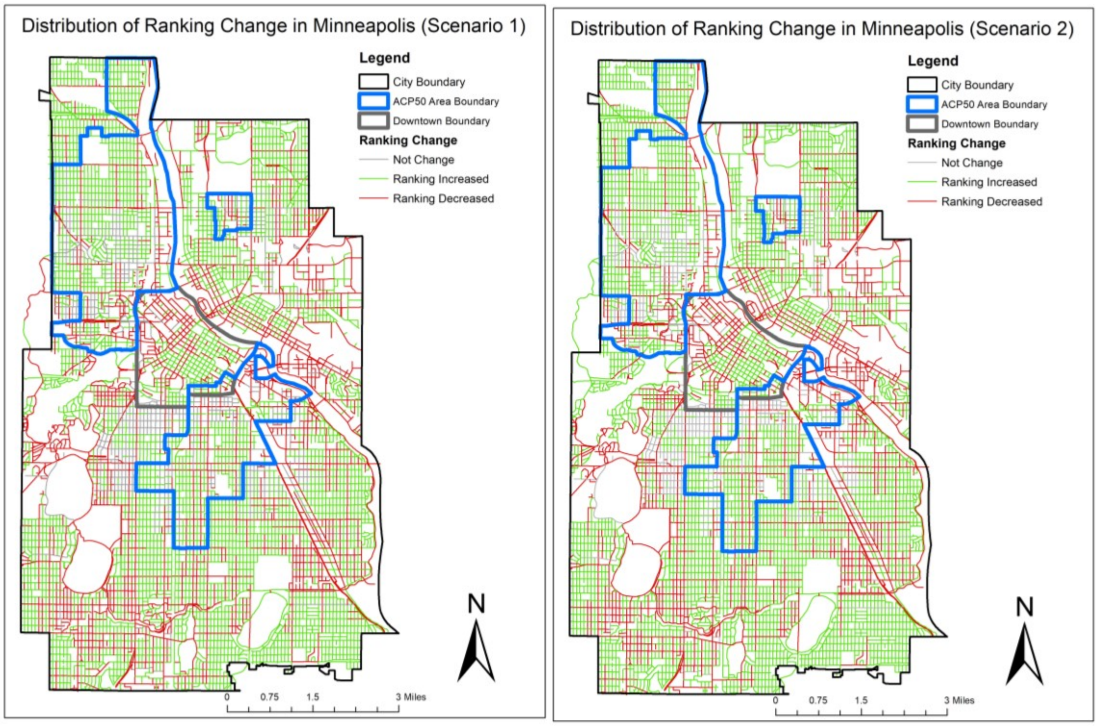
We follow ranking procedures used by DPW to illustrate the implications of augmenting the safety index. DPW uses street segments as the unit of analysis for street improvement prioritization. DPW first combines mid-blocks to obtain segments and then assigns scores, or point totals, to reflect the priority of individual segments. The DPW has identified 3,515 segments in Minneapolis.

We assigned point totals associated with our two new indices to the same segments identified by DPW and then compared the distributional outcomes, or rankings, to the DPW rankings based on its safety index. Specifically, we ranked segments using the two new indexes in place of the DPW safety index (Scenario 1) and by adding the two new indexes to the DPW index to create a three-dimensional index (Scenario 2). As expected, both Scenarios result different prioritization of segments. If the new indexes are used in place of the DPW safety index, 7.4% of all segments in the city retain their ranks, 57.2% have higher ranks, and 35.4% have lower ranks (Table 4.9). If the new indexes are used to create a three-dimensional index, 7.4% of segments retain their ranking, but 61.8% rise in the ranking, and 30.8% drop (Table 4.9). It is clear that use of different measures to address issues of safety (i.e., pedestrian and bicycle crash risk, and evenness in the distribution of risk) has the potential to change current rankings and prioritization based on safety.

**Table 4.9 Comparison between different applications of new indices**

| | Not change | Ranking increased | Ranking decreased |
|---|---|---|---|
| Safety (Scenario 0) v. Averaged predicted crash number + Crash risk Gini coefficient (Scenario 1) | 258 (7.4%) | 2,011 (57.2%) | 1,246 (35.4%) |
| Safety (Scenario 0) v. Safety + Averaged predicted crash number + Crash risk Gini coefficient (Scenario 2) | 258 (7.4%) | 2,174 (61.8%) | 1,083 (30.8%) |

To further illustrate the implications of changing the safety index used in ranking, we produced two visualizations, one to illustrate segments that change in rankings (Figure 4.7) and one to illustrate segments with the highest point totals (Figure 4.8). Neither visualization shows dramatic change. In Figure 4.7, the green segments, which are those that increase in ranking, appear to be predominantly local streets, while the red segments (i.e., those that drop in rankings) appear to include many higher functional class roads (i.e., arterials and collectors). This result indicates that that the principal change associated with both Scenarios 1 and 2 would be higher prioritization of local roads and lower prioritization of higher functional class roads.




(a) (b)

**Figure 4.7 Distribution of ranking change in Minneapolis for different scenarios**

Our second visualization (Figure 4.8) depicts segments with the highest scores in the DPW ranking and the two new scenarios. The highest score is defined slightly differently for different scenarios because Scenarios 1 and 2 involve multidimensional indexes. In the DPW ranking (Scenario 0), the high score, or priority point total, is 12. DPW assigned this score to about 33% of all segments scored. For Scenario 1, we assigned a point total of 20 as the threshold for the highest score. This threshold results in about 37% of all the segments ranked in highest point category. We assigned a threshold of 28 in Scenario 2 (i.e., the three dimensional index); use of this threshold also results in about 37% of all the segments in the highest point category. As shown in Figure 4.8, the variation in the spatial distribution of the highest ranked segments is relatively small. One noticeable difference is that fewer segments in the northern and southeastern peripheries of the city appear to be ranked highly in Scenarios 1 and 2. Because safety rankings account for only a fraction of total points used by DPW in the overall ranking of street improvement projects, additional research is required to determine if these new indexes would affect overall project prioritization.

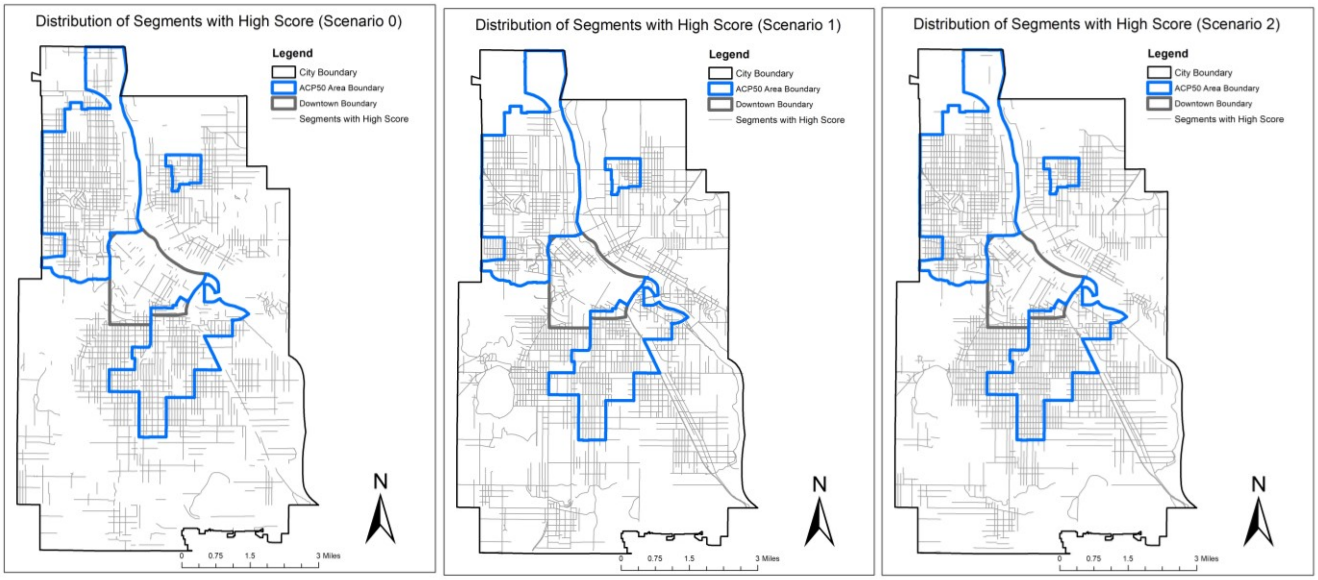



**Figure 4.8 Distribution of segments with high scores in Minneapolis for different scenarios**

# CHAPTER 5: CONCLUSIONS AND IMPLICATIONS

We estimated new models of pedestrian and bicycle crash risk at intersections and mid-blocks in Minneapolis, used these models to predict crashes at all intersections and mid-blocks in the city, and assessed the equity of distribution of crash risk. Our results show that crash risk, especially risk at intersections, is higher in neighborhoods that have lower incomes and higher populations of minorities than the rest of the city. We also developed new indices of crash risk and illustrated how findings and results could be used to inform ranking and prioritization of street improvement projects. Our crash models are distinctive relative to other models of crashes and crash risk reported in the literature (Table 2.1) in that they control simultaneously for vehicular, pedestrian, and bicycle exposure to risk across an entire street network in a major U.S. city. Our models show that, in most cases, crashes are correlated positively and significantly with exposure. Our crash models also confirm that different factors are associated with pedestrian and bicycle crashes and that these factors differ for both modes at intersections and mid-blocks. An implication of our results is that interventions and countermeasures to address crash risk need to be disaggregated and address simultaneously different risks faced by drivers, pedestrians, and bicyclists.

We assessed the equity of distribution of crash risk by using our crash models to predict pedestrian and bicycle crashes for all intersections and mid-blocks in the city and conducting significance tests between areas. A useful feature of our models is that by using the same factors to predict crashes for all locations, we obtain measures of risk at locations where no crashes historically have occurred. This feature helps to address a historical problem in crash analysis, specifically how to compare risk at locations where crashes have not occurred. We confirmed that crashes are distributed unevenly. Most intersections and mid-blocks have small crash risk, and crash risk is concentrated at a relatively small proportion of sites. We tested for significant differences in mean predicted crashes between racially-concentrated areas of poverty (APC50s) and other areas in the city, both inclusive and exclusive of the central business district (CBD). Mean predicted pedestrian and bicycle crashes at intersections were substantially higher in APC50s than in non-APC50 areas. These differences increased substantially for pedestrian and bicycle crashes when the CBD was excluded from the analyses. These analyses of crash risk corroborated findings reported by the city based solely on locations of existing crashes and affirm efforts to address inequities in crash risk.

To illustrate how increased emphases can be placed on pedestrian and bicycle crashes in the prioritization of street improvement projects, we used our predicted crash estimates to create a combined pedestrian and bicycle crash index for all mid-blocks in the city. We also used our predicted crashes to calculate Gini Coefficients at the census block level to obtain another index of unevenness in the distribution of crash risk. We then showed how these two indices could be used to complement or augment rankings currently used by the city to prioritize street improvement projects that are based on crash rates calculated using all types of crashes, not only pedestrian and bicycle crashes. The results, not surprisingly, showed that introduction of new criteria for ranking, specifically new measures of pedestrian and bicycle crash risk, could change the order of ranking and therefore project prioritization.

Our research has a number of limitations that can be addressed over time through additional analyses and as more data are collected. Our crash models are based on a limited number of locations where we have both crash records and estimates of exposure for vehicles, pedestrian, and cyclists. Over time, as Minneapolis expands its pedestrian and bicycle monitoring program, we can obtain a larger sample and re-estimate our crash models. Another limitation of our models is that the time periods for our crash data, our estimates of exposure, and our other independent variables are different. Most important, we use estimates of peak-hour, weekday, summer and fall exposure in our models, while our crash data includes crashes throughout the year, for all days of the week, and for all times of day. Our rationale for using these measures of pedestrian and bicycle exposure is that we believe the distributions of peak-hour and total are similar. Nevertheless, this temporal mismatch means our results must be interpreted with caution. Another limitation of our pedestrian mid-block model is that our estimate of pedestrian exposure is of pedestrians on the sidewalk, not pedestrians actually in the street or crossing at mid-block. We believe these two measures also are likely to be correlated, but field studies would be required to confirm this hypothesis. More generally, new data collection initiatives could help to address each of these limitations. Another approach to addressing these limitations is to conduct scenario and simulation analyses based on subsets of our data to assess the stability of these results. These types of analyses are beyond the scope of this study. We also believe our models can be strengthened by refining the sets of variables within them and by addressing potential issues such as spatial autocorrelation. As noted, our models of pedestrian and bicycle crash risk at both intersections and mid-blocks generally include the same variables. We retained variables that proved to be insignificant both for theoretical reasons and to illustrate differences in factors associated with different types of crashes at different locations. Refinement of these models potentially can add to findings.

With respect to the implications of our research for ranking, our results show that the addition of additional criteria to existing safety-related criteria has the potential to change rankings. As we have already noted, this finding in and of itself is not surprising. The important insight from this finding is that viable methods of increasing emphases on pedestrian and bicycle safety exist and potentially could be incorporated into rankings. A limitation of our work is that we did not simulate effects of our new indices on overall project ranking. The Minneapolis DPW currently uses an array of factors such as street pavement condition in its ranking system; safety and equity are only two of the factors. Our analyses are limited to a marginal assessment, primarily because of time limitations for the study. In addition, in our illustration of the effects of using our pedestrian and bicycle crash indexes in ranking, we follow DPW procedures and use four categories for assignment points. Different methods of assigning points that better reflect the distribution of crash risk are available and potentially could be incorporated. Future studies also can address this avenue of research.

Last, our emphases in this report have been on technical methods of assessing crash risk with the goal of informing project prioritization. All measures of crash risk are limited and imperfect and therefore entail uncertainty. These facts mean that measures of crash risk incorporated in project ranking systems introduce error and uncertainty that, potentially, can confound rankings. The choice of which factors to include in project ranking is a professional, subjective choice, informed both by technical analyses and community values. In response to feedback from community outreach, Minneapolis is seeking strategies to make its rankings place additional emphases on safety and equity, among other factors. One approach to assessing the desirability of measures like those developed in this research could involve collaboration with community residents to obtain additional, specific feedback on measures and procedures used for ranking.

# REFERENCES


Carlson, K., Murphy, B., Ermagun, A., Levinson, D. M., & Owen, A. (2018). *Safety in numbers: Pedestrian and bicyclist activity and safety in minneapolis.* Minneapolis: City of Minneapolis.

Chen, P. (2015). Built environment factors in explaining the automobile-involved bicycle crash frequencies: A spatial statistic approach. *Saf. Sci., 79*, 336–343. doi.org/10.1016/j.ssci.2015.06.016

City of Minneapolis. (2017a). Annual update 2017. 20-Year Streets Funding Plan. Minneapolis: City of Minneapolis.

City of Minneapolis. (2017b). *2017 City of Minneapolis pedestrian crash study*. Minneapolis: City of Minneapolis.

Cottrill, C. D., & Thakuriah, P. V. (2010). Evaluating pedestrian crashes in areas with high low-income or minority populations. *Accid. Anal. Prev., 42*, 1718–1728.

Daniels, S., Brijs, T., Nuyts, E., Wets, G. (2009). Injury crashes with bicyclists at roundabouts: Influence of some location characteristics and the design of cycle facilities. *J. Safety Res. 40*, 141–148. doi.org/10.1016/j.jsr.2009.02.004

Department of Public Safety. (2018). Crashes in Minnesota from 2005 to 2017. St. Paul: Minnesota Department of Public Safety.

Dumbaugh, E., & Li, W. (2011). Designing for the safety of pedestrians, cyclists, and motorists in urban environments. *J. Am. Plan. Assoc., 77*, 69–88. doi.org/10.1080/01944363.2011.536101

El-Basyouny, K., & Sayed, T. (2013). Safety performance functions using traffic conflicts. *Saf. Sci.. 51*, 160–164. doi.org/10.1016/j.ssci.2012.04.015

Elvik, R. (2013). Can a safety-in-numbers effect and a hazard-in-numbers effect co-exist in the same data? *Accid. Anal. Prev., 60*, 57–63. doi.org/10.1016/j.aap.2013.08.010

Elvik, R. (2009). The non-linearity of risk and the promotion of environmentally sustainable transport. *Accid. Anal. Prev., 41*, 849–855. doi.org/10.1016/j.aap.2009.04.009

Elvik, R., & Bjørnskau, T. (2017). Safety-in-numbers: A systematic review and meta-analysis of evidence. *Saf. Sci., 92*, 274–282.

Federal Highway Administration (2018a). *Guide for scalabe risk assessment methods for pedestrians and bicyclists.* Washington, DC: Federal Highway Administration.

Federal Highway Administration (2018b). Systemic pedestrian safety analysis. Retrieved from https://doi.org/10.17226/25255

Gladhill, K., Monsere, C. (2012). Exploring traffic safety and urban form in Portland, Oregon. *Transp. Res. Rec. J. Transp. Res. Board, 2318*, 63–74. doi.org/10.3141/2318-08

Hankey, S., & Lindsey, G. (2016). Facility-demand models of peak period pedestrian and bicycle traffic. *Transp. Res. Rec. J. Transp. Res. Board, 2586*, 48–58. doi.org/10.3141/2586-06

Jacobsen, P. (2003). Safety in numbers: More walkers and bicyclists, safe walking and biking. *Inj Prev, 9*, 205–209.

Jonsson, T. (2005). Predictive models for accidents on urban links—A focus on vulnerable road users. *Bull. Inst. Technol. Dep. Technol. Soc*, *226*, 1-142.

Kim, S., Ulfarsson, G. F., & Todd Hennessy, J. (2007). Analysis of light rail rider travel behavior: Impacts of individual, built environment, and crime characteristics on transit access. *Transp. Res. Part A Policy Pract., 41*, 511–522. doi.org/10.1016/j.tra.2006.11.001

Lee, C., & Abdel-Aty, M. (2005) Comprehensive analysis of vehicle-pedestrian crashes at intersections in Florida. *Accid. Anal. Prev., 37*, 775–786. doi.org/10.1016/j.aap.2005.03.019

Lindsey, G., Wang, J., Hankey, S., & Pterka, M. (2018). Modeling bicyclist exposure to risk and crash risk: Some exploratory studies. University of Minnesota, Roadway Safety Institute, CTS 18-15.

Loukaitou-Sideris, A., Liggett, R., & Sung, H. G. (2007). Death on the crosswalk: A study of pedestrian-automobile collisions in Los Angeles. *J. Plan. Educ. Res., 26*, 338–351. doi.org/10.1177/0739456X06297008

Maciag, M. (2014). Pedestrians dying at disproportionate rates in America's poorer neighborhoods. Retrieved from http://www.governing.com/topics/public-justice-safety/gov-pedestrian-deaths-analysis.html (accessed 2.21.19).

Nantulya, V. M., & Reich, M. R. (2003). Equity dimensions of road traffic injuries in low- and middle-income countries. *Inj. Control Saf. Promot., 10*, 13–20. doi.org/10.1076/icsp.10.1.13.14116

National Academies of Science Engineering and Medicine. (2008). *Pedestrian safety prediction methodology.* Washington, DC: The National Academies Press. Retrieved from https://doi.org/10.17226/23083

Nordback, K., Marshall, W. E., & Janson, B. N. (2014). Bicyclist safety performance functions for a U.S. city. *Accid. Anal. Prev., 65*, 114–122. doi.org/10.1016/j.aap.2013.12.016

Park, J., Abdel-Aty, M., Lee, J., & 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. *Accid. Anal. Prev., 74*, 179–191. doi.org/10.1016/j.aap.2014.10.024

Robinson, D. (2005). Safety in numbers in Australia: More walkers and bicyclists, safer walking and cycling. *Heal. Promot. J. Aust., 16*, 47–51.

Schepers, J. P., Kroeze, P. A., Sweers, W., & Wüst, J. C. (2011). Road factors and bicycle-motor vehicle crashes at unsignalized priority intersections. *Accid. Anal. Prev., 43*, 853–861. doi.org/10.1016/j.aap.2010.11.005

Schneider, R., Diogenes, M., Arnold, L., Attaset, V., Griswold, J., & Ragland, D. (2010). Association between roadway intersection characteristics and pedestrian crash risk in Alameda County, California. *Transp. Res. Rec. J. Transp. Res. Board, 2198*, 41–51. doi.org/10.3141/2198-06

Siddiqui, C., Abdel-Aty, M., & Choi, K. (2014). Implications of pedestrian safety planning factors in areas with minority and low-income populations. *Int. J. Sustain. Transp., 8,* 360–381.

doi.org/10.1080/15568318.2012.702853

Siddiqui, C., Abdel-Aty, M., & Choi, K. (2012). Macroscopic spatial analysis of pedestrian and bicycle crashes. *Accid. Anal. Prev., 45*, 382–391. doi.org/10.1016/j.aap.2011.08.003

Strauss, J., Miranda-Moreno, L. F., & Morency, P. (2013). Cyclist activity and injury risk analysis at signalized intersections: A Bayesian modelling approach. *Accid. Anal. Prev., 59*, 9–17. doi.org/10.1016/j.aap.2013.04.037

Thomas, L., Lan, B., Sanders, R. L., Frackelton, A., Gardner, S., & Hintze, M. (2017). Changing the future?: Development and application of pedestrian safety performance functions to prioritize locations in Seattle, Washington. *Transp. Res. Rec. J. Transp. Res. Board, 2659*, 212–223. doi.org/10.3141/2659-23

Turner, S., Wood, G., Hughes, T., & Singh, R. (2011). Safety performance functions for bicycle crashes in New Zealand and Australia. T*ransp. Res. Rec. J. Transp. Res. Board, 2236*, 66–73. doi.org/10.3141/2236-08

Ukkusuri, S., Miranda-Moreno, L. F., Ramadurai, G., & Isa-Tavarez, J. (2012). The role of built environment on pedestrian crash frequency. *Saf. Sci., 50*, 1141–1151. doi.org/10.1016/J.SSCI.2011.09.012

Wang, J., & Lindsey, G. (2017). Equity of bikeway distribution in Minneapolis, Minnesota. *Transp. Res. Rec. J. Transp. Res. Board*, 2605, 18–31. doi.org/10.3141/2605-02

Wei, V. F., & Lovegrove, G. (2012). Sustainable road safety: A new (?) neighborhood road pattern that saves VRU lives. *Accid. Anal. Prev., 44*, 140–148. doi.org/10.1016/j.aap.2010.12.005

Yasmin, S., & Eluru, N. (2016). Latent segmentation based count models: Analysis of bicycle safety in Montreal and Toronto. *Accid. Anal. Prev., 95*, 157–171. doi.org/10.1016/j.aap.2016.07.015

Zegeer, C., Lyon, C., Srinivasan, R., Persaud, B., Lan, B., Smith, S.....& Sundstrom, C. (2017). Development of crash modification factors for uncontrolled pedestrian crossing treatments. *Transp. Res. Rec. J. Transp. Res. Board, 2636*, 1–8. doi.org/10.3141/2636-01