Safety in Numbers and Safety in Congestion for Bicyclists and Motorists at Urban Intersections

Kristin Carlson
Graduate Research Assistant
University of Minnesota, Department of Civil, Environmental, and Geo- Engineering
500 Pillsbury Drive SE, Minneapolis, MN 55455 USA
carl4498@umn.edu

Alireza Ermagun *(Corresponding Author)*
Postdoctoral Fellow
Northwestern University, Department of Civil and Environmental Engineering
Technological Institute, 2145 Sheridan Rd, Evanston, IL 602088 USA
alireza.ermagun@northwestern.edu

Brendan Murphy
Research Fellow
University of Minnesota, Center for Transportation Studies
200 Transportation and Safety Building, 511 Washington Ave SE, Minneapolis, MN 55455 USA
murph677@umn.edu

Andrew Owen
Director of Accessibility Observatory
University of Minnesota, Center for Transportation Studies
200 Transportation and Safety Building, 511 Washington Ave SE, Minneapolis, MN 55455 USA
aowen@umn.edu

David Levinson
Professor
University of Sydney, School of Civil Engineering
david.levinson@sydney.edu.au

Paper submitted for:
Presentation at 97thAnnual Transportation Research Board Meeting, January 2018
Standing Committee on

words + 4 figures + 2 tables
July 30, 2017

ABSTRACT

This study assesses the estimated crashes per bicyclist and per vehicle as a function of bicyclist and vehicle traffic, and tests whether greater traffic reduces the per-car crash rate. We present a framework for comprehensive bicyclist risk assessment modeling, using estimated bicyclist flow per intersection, observed vehicle flow, and crash records. Using a two-part model of crashes, we reveal that both the annual average daily traffic and daily bicyclist traffic have a diminishing return to scale in crashes. This accentuates the positive role of safety in numbers. Increasing the number of vehicles and cyclists decelerates not only the probability of crashes, but the number of crashes as well. Measuring the elasticity of the variables, it is found that a 1% increase in the annual average daily motor vehicle traffic increases the probability of crashes by 0.14% and the number of crashes by 0.80%. However, a 1% increase in the average daily bicyclist traffic increases the probability of crashes by 0.09% and the number of crashes by 0.50%. The saturation point of the safety in numbers for bicyclists is notably less than for motor vehicles. Extracting the vertex point of the parabola functions examines that the number of crashes starts decreasing when daily vehicle and bicyclist traffic per intersection exceed 29,568 and 1,532, respectively.

Keywords: Safety; Bicyclist crashes; Returns to scale; Road intersection

INTRODUCTION

While walking and bicycling have been shown to be positively correlated with curbing air pollution and promoting health, over half of the global annually reported 1.25 million vehicle crashes involve a pedestrian, bicyclist, or motorcyclist. The number of bicyclists compared to motorized vehicles would suggest nearly negligible annual traffic crashes, yet bicyclists make up 2% of all traffic related deaths (1). Active modes of travel (e.g. walking and bicycling) as a set of modes tends to be less safe than motor vehicles on a per kilometer basis (2). This holds true in most average developed urban areas, except where specific programs and treatments have been employed to address the safety concerns (3).

The term "safety in numbers" (SIN) was coined in 1949, when Smeed (4) showed that road fatalities per vehicle were lower in countries with more driving. He demonstrated that an exponential curve describes the relationship between fatal vehicular crashes and vehicle kilometers traveled (VKT). SIN refers to the phenomenon that bicyclists as road users become safer when there are more riders present in a given locale or area. Much of the previous research has echoed his finding and corroborated the existence of SIN effect. A variety of methodologies were employed to try to capture the magnitude and the contributing factors to the SIN effect, while controlling for environment and human behavior. One such study took place in Hamilton, Ontario, Canada where pedestrian flow was compared to the crash rates (5). Data collected from 300 signalized intersections from 1983-1986 contained pedestrian crashes and estimated pedestrian and vehicular flows. Decreasing per pedestrian risk was associated with increasing pedestrian flows. Conversely, increasing vehicle flows was associated with increased pedestrian risk. The crash counts at each intersection were considered as a Poisson random variable. This study found that drivers seem to expect pedestrians when the pedestrian flow is over 30 pedestrians per hour. It was also found that the level of bicyclist flow is more important for bicyclist safety than the level of vehicular exposure. A similar study was conducted in Sweden in 1996, which compared bicyclist counts against crashes at 95 intersections. Once again, an inverse relationship was found between bicyclist counts and the number of bicyclist-auto crashes (6).

A study in 2003 used five data sets, which included three population level and two time series datasets. It was found that the SIN effect is "consistent across communities of varying size, from specific intersections to cities and countries, and across time periods" (7). This study used a dataset that linked the number of crashes with the amount of walking and bicycling, however vehicle flow was not an explanatory variable. The model was estimated as a power curve. It was found that the number of pedestrians and bicyclists struck by vehicle vary by the 0.4 power of the pedestrian or bicyclist traffic. Earlier, researchers in Australia had tested the power model on a dataset that contained over 100 years of crash information (8). Another more recent Australian study (9) used three types of pedestrian or bicyclist injury datasets to recreate the negative exponential curve. Safety in numbers was found to exist in Australia with a similar exponential relationship compared to the American studies. If cycling doubles, the risk per kilometer falls by about 34%.

After reviewing several years of studies that were conducted around the world to verify the SIN effect, Elvik (9) found that transferring trips from motorized vehicles to walking and biking reduces the number of crashes. This study changed functional form based on the type of crash involving a pedestrian or bicyclist (multi-vehicle, single-vehicle). The parameters that were varied included number of motor vehicles, pedestrians, bicyclists, and the coefficient values for pedestrian and bicyclist crashes. The exponential form was used and the risk calculated for different Annual Average Daily Traffic (AADT) values (2,000-30,000). It was found that, in theory, the total number of crashes could go down if a substantial share of trips by motorist transport is transferred to walking or cycling (9).

The contribution of the current research to the literature is twofold:

- SIN is well-supported by bicyclist crash data across a number of studies in various environments (5, 7, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19). The most frequently cited hypothesized cause of the SIN effect is that motorists adapt their driving behavior when traversing roadway that frequently carries pedestrian and bicyclist traffic (7, 20). The SIN concept has seen relatively widespread adoption in urban planning schools of thought, though its temporal causality is not clear-cut (10), and it is commonly discussed only in the context of bicyclist risk depending on bicyclist flow levels. Little is known, however, about the Safety in Congestion (SIC) effect. We hypothesize that greater traffic flow reduces the per-car crash rate. The reasoning is that greater congestion reduces vehicle speeds thereby giving drivers greater reaction time to reduce the severity of a crash or avoid one altogether. In particular, we assess the estimated crashes per bicyclist and per car outcomes as a function of bicyclist and vehicle traffic. This supports SIC effect by either modes of transport.
- Much of the previous research has focused on aggregate data, typically at the level of Transport Analysis Zones (TAZs), which is too coarse to allow robust analysis of non-motorized travel (21, 22). Regional Travel Surveys consider many trip purposes, but are similarly coarse, and typically have sample sizes too small to allow for robust city-to-city comparison. We present a framework for comprehensive bicyclist risk assessment modeling, using estimated bicyclist flow per intersection, observed vehicle counts, and crash records. The motivation for using models of bicyclist traffic is in supplementing the sparse data currently available, in order to assess bicyclist risk-burdens of collisions at every intersection in Minneapolis, Minnesota. Bicyclist risk-burdens - the risk of an individual bicyclist being struck by a vehicle - are calculated and compared for both the raw and predicted crash per bicyclist data sets. This process allows us to construct a more complete spatial picture of how bicyclist collision risk varies throughout an urban area at the level of individual intersections, based on data widely available to practitioners, transportation authorities, and the public.

The remainder of this study is organized as follows. First, we introduce the data used in this research along with the data extraction and preparation process. Following the discussion of the methodology, we provide the results of the modeling procedure along with an in-depth interpretation of interest variables both qualitatively and quantitatively. We conclude the paper with summarizing the key findings and opening new research avenues.

COMPREHENSIVE BICYCLIST RISK ASSESSMENT FRAMEWORK

Although proper placement of bicyclist treatments and improvements has implications to both safety (21) and accessibility and mode choice (22), proper information regarding estimated non-motorized traffic levels is needed to locate areas in need of improvement. In determining salient locations for non-motorized improvements, it is important to have accurate records of both existing and potential travel demand (e.g. current levels of biking in a neighborhood, as well as good models of increased demand due to potential treatments); however good quality, high-granularity datasets for non-motorized travel can be difficult to obtain, especially standardized for national spatial inventories (23). Hence, planners and advocates must frequently rely on estimation models for non-motorized traffic, and various methods can suffer from issues of data quality, granularity, and the presence of location-specific variables (24).

Many of the issues with the collection of standardized non-motorized transportation data have to do with the factors that influence pedestrian and bicyclist behavior. A model of active transport risk assessment is uninformative if the bicyclist and vehicular flows do not accurately represent corresponding levels *in situ*, and many cities do not have dense data sets of active transport flow levels, instead favoring counts of vehicle traffic. As such, active transport flow levels must be extrapolated from sparse datasets using comprehensive methodologies. Population and employment data are well-documented by the US Census Bureau to the Census Block level of resolution, and general socioeconomic characteristics are maintained as well, and can have significant influence (21). However, more specific socioeconomic characteristics are salient in non-motorized travel beyond just adjusted income levels, as well as weather variables (25) and latent, subjective variables such as visibility and perceptions of lighting, which can be more difficult to obtain at high spatial resolution (26), and can complicate inter-city comparisons. For these reasons, as well as the overall lack in non-motorized travel counts for many communities, methods of estimating pedestrian and bicyclist behavior that do not rely heavily on high-resolution count data are applied in this study.

We employed two types of data to create the crash prediction models: (1) Estimated Bicyclist Activity Data and (2) Bicyclist-Auto Crash Prediction Model Data. The former is borrowed from a previously developed land use and transportation network regression model. The latter is raw observations taken from local records.

Estimated Bicyclist Activity Data

A reduced-form core facility demand model gets around the issue of data quality and granularity by using easily retrievable data sets to predict bicyclist counts at the intersection level (27). This model was developed using existing pedestrian and bicyclist counts, land use variables, and transportation network variables extracted from the State of Minnesota. The facility demand model estimates peak-period traffic flow at intersections where counts are unavailable. The outcome was an estimated comprehensive pedestrian and bicyclist count dataset for the City of Minneapolis, which can be used to examine trends related to bicyclist activity.

Facility-demand bicyclist estimates developed by Hankey and Lindsey (27) are derived from a reduced-form core model, which allows practitioners to use easily retrievable data sets to predict bicyclist counts at the intersection level. The model used to develop the estimated bicyclist traffic was derived from counts taken in September during the peak-period (4:00 PM - 6:00 PM).

Independent variables were selected based on their known likelihood to affect a citizen's propensity to bike. The 2014 employment accessibility for the Twin Cities region was included (28) along with land use variables (i.e., industrial area, population density, retail area, open space) and the number of bicyclist facilities. Temporal variables such as temperature and precipitation were used to account for the weather shifts in Minnesota and the resulting bicyclist activity. The 2010 U.S. Census core-based statistical areas for Minneapolis - St. Paul were used to cordon bicyclist facility counts, population, and demographic information. For more information on the facility-demand modeling procedure and detailed discussion, we refer readers to (27). For a similar analysis using a different methodology see (29), while for a similar analysis of the pedestrian safety in numbers effect see (30). We list the core datasets to estimate bicyclist activity as follows:

- PM peak period bicyclist counts observed in September from 2007-2015, conducted by the City of Minneapolis Department of Public Works (DPW) and Transit for Livable Communities (TLC)
- Employment accessibility within 5-60 minutes of walking in 2014, University of Minnesota Accessibility Observatory (**?**)
- Land use statistics, Metropolitan Council 2015
- U.S. Census TIGER 2010 datasets: blocks, core-based statistical area (CBSA) for Minneapolis-St. Paul
- Tabulation of yearly bicyclist facilities, (DPW) and (TLC) 2007-2015
- Weather parameters, (DPW) and (TLC) 2007-2015

The estimated counts were chosen for this study to expand the dataset available for analysis. Figure 1 depicts the number of daily bicyclists observed to pass through each intersection shown. In comparison, Figure 2 shows the estimated bicyclist counts that were made available through the facility-demand model.

Bicyclist-Auto Crash Prediction Model Data

The crash prediction model presented in this report uses four independent variables to predict the number of crashes at a given intersection in Minneapolis over a 14 year period. The independent variables include the estimated bicyclist Turning movement counts (TMC) and the observed AADT and their quadratic forms for 489 intersection in Minneapolis. For comparative purposes, raw bicyclist TMC were assessed and plotted to verify that using estimated bicyclist TMCs would be an improvement compared with the raw data. The number of bicyclist-auto crashes that were recorded from 2000 to 2013 at each of these intersections was included as the dependent variable. A 2014 OpenStreetMap extract of the Twin Cities region was used to geocode crash records to intersections. OpenStreetMap is an open-source platform of free and reusable geospatial data. We list the core datasets to estimate bicyclist-auto crashes follows:

- OpenStreetMap (OSM) North America extract, retrieved July 2016

- Raw bicyclist Turning Movement Counts (TMC) 2007-2014, City of Minneapolis
- Estimated bicyclist Turning Movement Counts September 4-6 PM, City of Minneapolis
- Annual Average Daily Traffic (AADT) measurements 2000-2013, City of Minneapolis
- Traffic crash records 2000-2013, City of Minneapolis

DATA PREPARATION

Intersection locations were determined from OSM road centerline data for the Minneapolis - St. Paul Core-Based Statistical Area (CBSA). To get a sense for the magnitude of bicyclist traffic throughout Minneapolis, Figure 3 was developed to visualize the distribution of bicyclist activity. The estimated bicyclist TMC was geocoded to intersections for a single value of bicyclist traffic at each intersection from 4:00 PM - 6:00 PM. Estimated peak-hour bicyclist flows were expanded to 24-hour counts using bicyclist traffic count factors to extrapolate the estimated counts (31). The AADT records from 2000 to 2013 were averaged over those years and assigned to intersections by applying a mid-block buffer around each intersection in QGIS and summing the cross streets for a single value of vehicle traffic. Crash records were geocoded to intersections using the OSM extract and QGIS. Figure 3 shows the locations and levels of bicyclist-auto crashes that occurred at each intersection in the test dataset. GIS work was performed in QGIS and PostGIS; statistical work done in Stata and NLOGIT software. Table 1 gives the description and statistics of the variables used in this study for both parts of the modeling procedure.

TABLE 1 : Bike activity and safety dataset summary statistics

Variable	Description	Min	Max	Mean	Std. Dev.
	Training Data (n=383)				
Crashes	Cumulative crashes from 2000-2013	0	16	1.50	2.4
Vehicle Traffic	Mean daily traffic per intersection 2000-2013	252	30798	8584.9	5693.0
Bicyclist Traffic	24 hour bicyclist count per intersection	37	3,935	793.1	562.6
	Test Data (n=106)				
Crashes	Cumulative crashes from 2000-2013	0	11	1.54	2.3
Vehicle Traffic	Mean daily traffic per intersection 2000-2013	440	25927	8424.8	4920.1
Bicyclist Traffic	24 hour bicyclist count per intersection	57	2,163	877.2	539.7

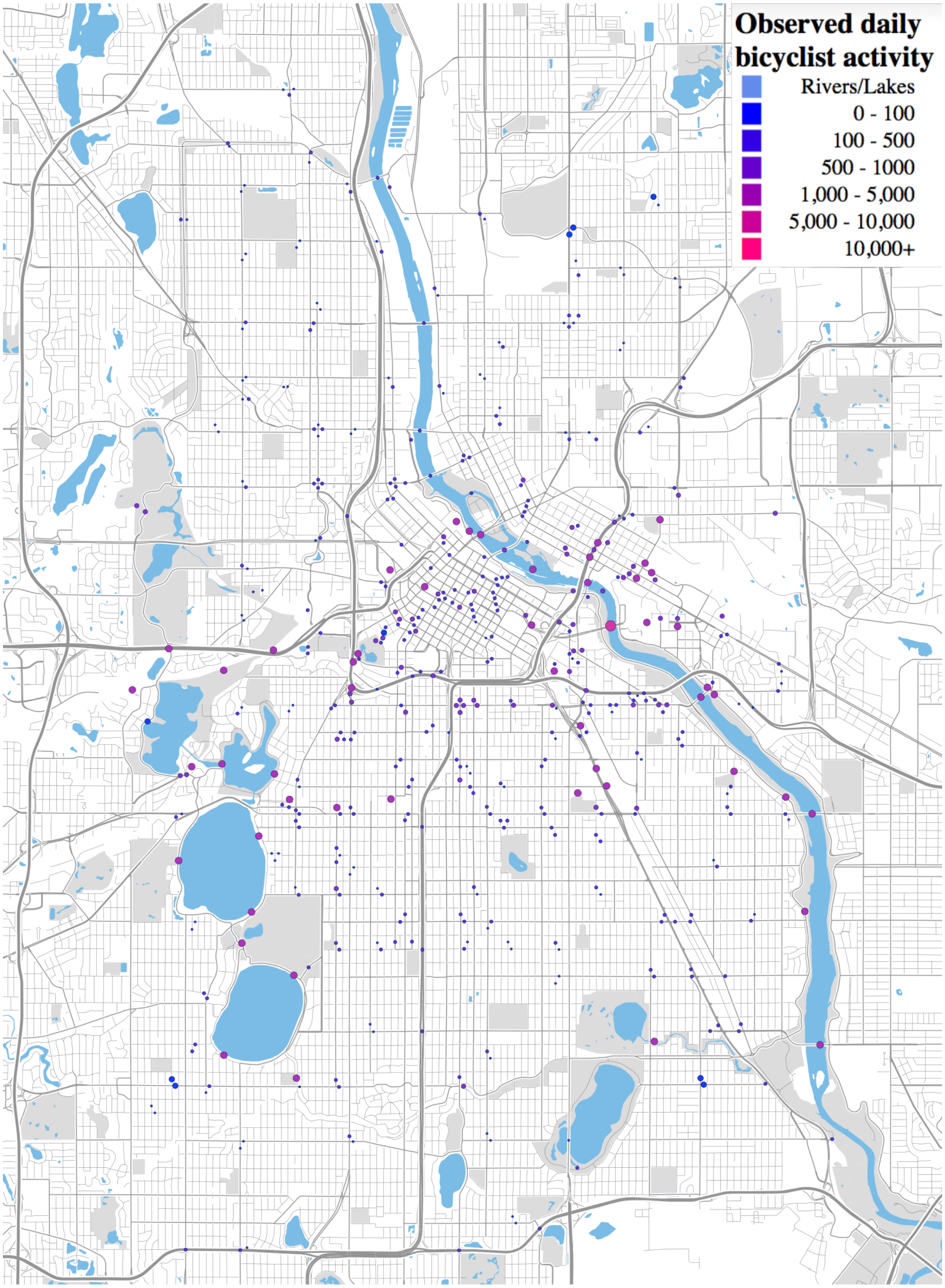


FIGURE 1 : Observed levels of daily bicyclist activity in Minneapolis, 2007-2014

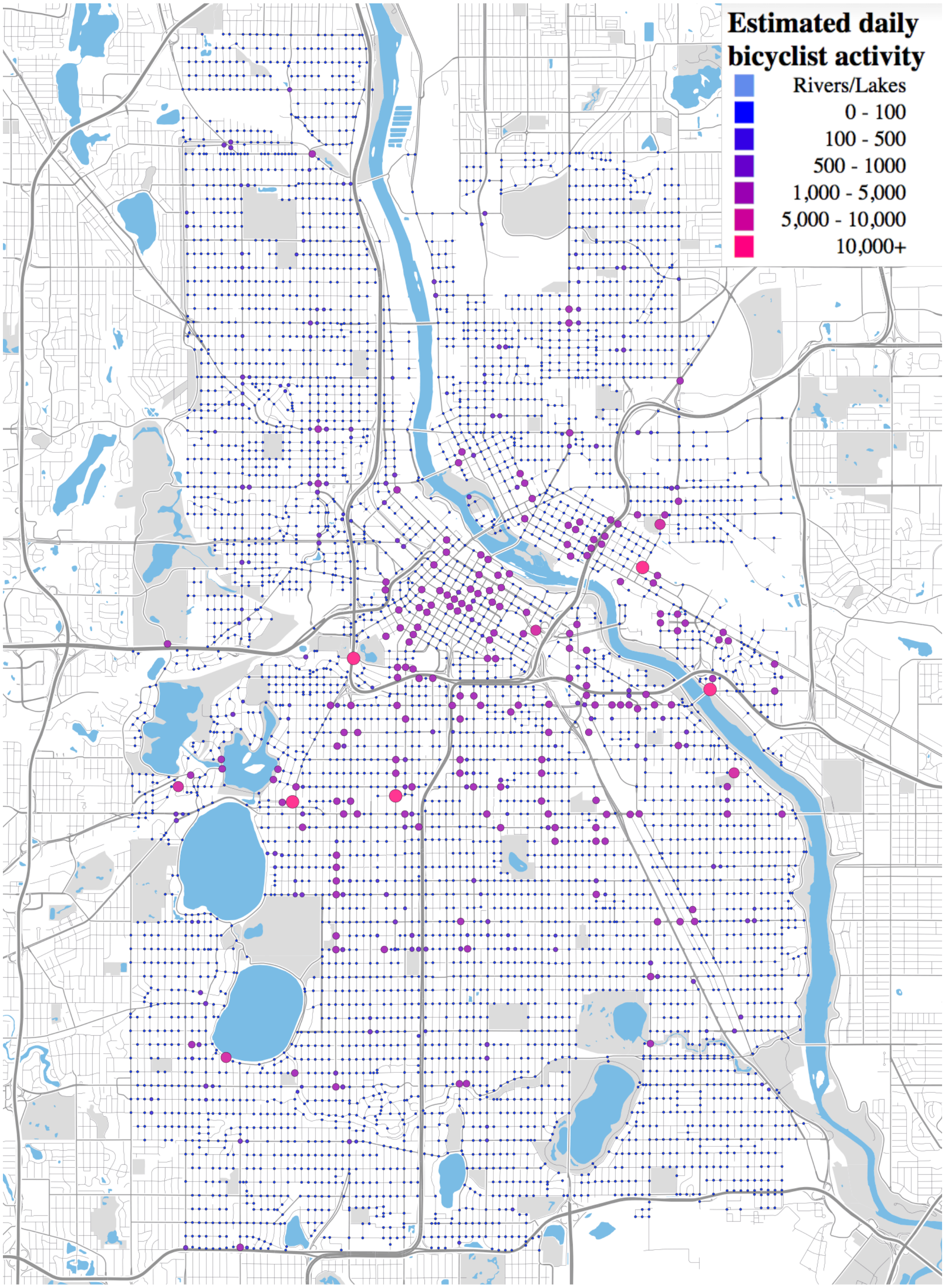


FIGURE 2 : Estimated levels of daily bicyclist activity in Minneapolis.

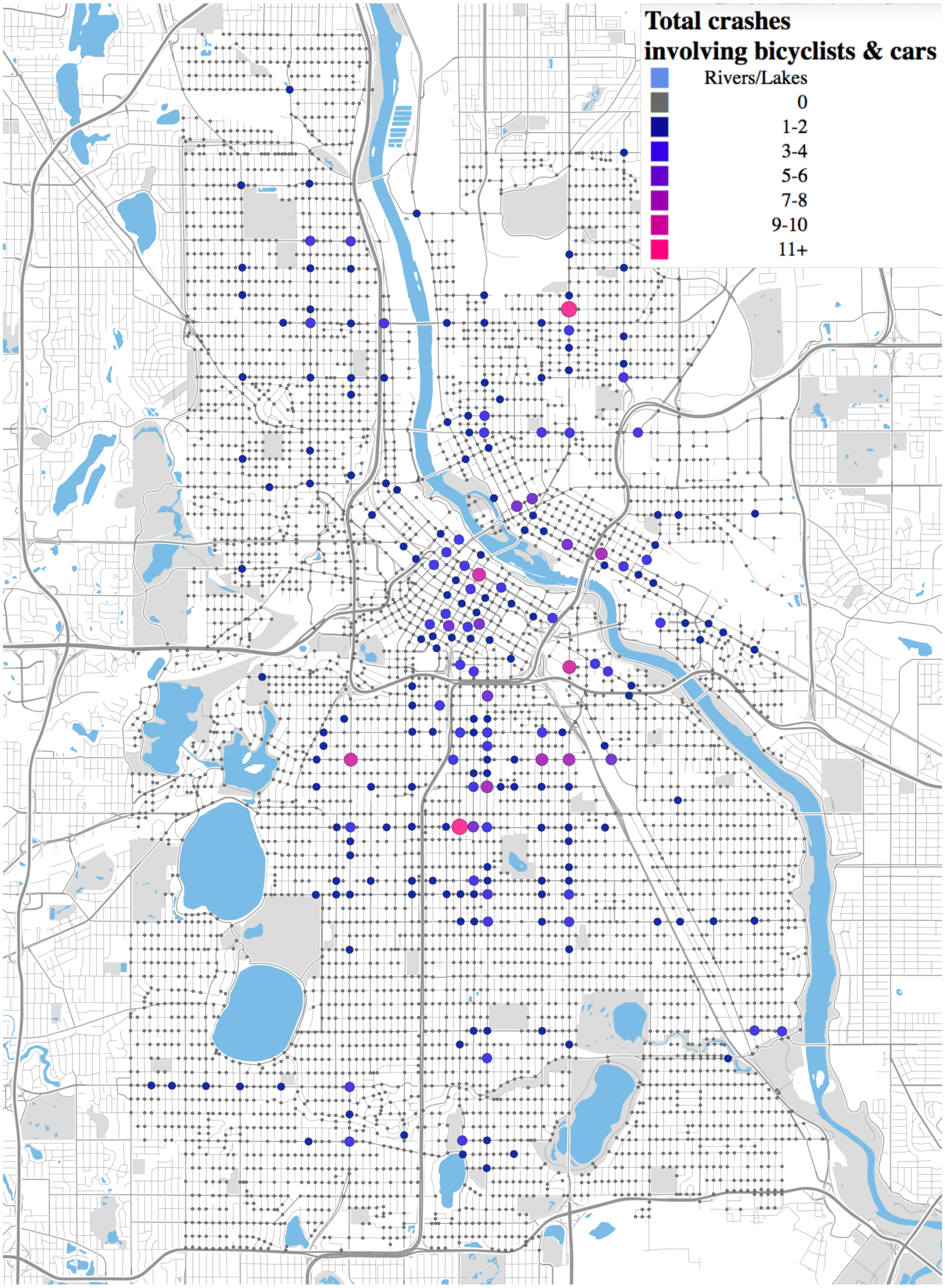


FIGURE 3 : Raw levels of bicyclist-auto crashes in Minneapolis, 2000-2013.

Methodology: Two-part Model of Crashes

Modeling the number of bicyclist-auto crashes requires a different approach than traditional count outcomes or linear models, as the majority of intersections have no cumulative crashes over time. A high proportion of zeros in the distribution of the number of crashes variable means that standard approaches such as least squares regression misrepresents the results. To overcome this challenge, statisticians have introduced methods that account for such infrequent distributions including the Heckit, latent Heckit, and two-part methods. These models are generally applied to model distributions of continuous and non-negative data, which contain a large proportion of zero observations. We refer the reader to Dow and Norton (32) for a broad discussion over the advantages and disadvantages of each method.

The descriptive analysis demonstrates that nearly half of the intersections used in this study have no reported crashes between 2000 and 2013. To represent the marginal effects of exogenous variables accurately, we employ a two-part model of crashes. This model comprises Probit regression for the first-part and Poisson regression for the second-part of the model. The former predicts the probability of zero versus non-zero crashes at a given intersection, and the latter model sites with one or more than one crash. It was assumed that the crash data were not over-dispersed and the excess zeros assumption was addressed in the first-part of the modeling procedure.

Student's t-statistic and Adjusted Pseudo R^2 are measured to test the statistical significance of variables and the general fit of the models, respectively. The data is randomly divided into two portions: An 80% part for training the models and a 20% part for testing the prediction power of the models. We postulate two main hypotheses:

- The number of crashes has a diminishing return to scale with respect to annual average daily motor vehicle traffic.
- The number of crashes has a diminishing return to scale with respect to daily bicyclist traffic.

To test these hypotheses, we embed the annual average daily motor vehicle traffic and daily bicyclist traffic (DBT) along with their quadratic form in the models. Doing so enables us to capture the more realistic association between number of crashes and the interest variables. The modeling results are outlined in Table 2.

TABLE 2 : Two-part model results—bike crashes

Variable	Part One		Part Two	
	Y_1 No Crash: 0, Crash: 1		Y_2 Number of Crashes	
Description	*Coefficient*	*t-test*	*Coefficient*	*t-test*
Vehicle Traffic	1.05×10^{-4}	3.03	9.58×10^{-5}	3.70
(Vehicle Traffic)^2	-3.82×10^{-9}	-2.74	-1.62×10^{-9}	-1.66
Bicyclist Traffic	6.65×10^{-4}	2.12	9.90×10^{-4}	3.61
(Bicyclist Traffic)^2	-2.53×10^{-7}	-1.91	-3.23×10^{-7}	-2.57
Constant	-0.80	-3.57	-0.14	-0.75
Number of observations	383		190	
Pseudo R^2	0.03		0.352	

GENERAL DISCUSSION

Evidence of Safety in Numbers

The student's t-statistic measurement indicates that all variables are significant at the 90% confidence interval. Looking at the first-part of the model, the downward parabola form of the *Vehicle Traffic* variable demonstrates that the probability of crashes has a diminishing return to scale considering the annual average daily traffic as an input. This confirms our hypothesis that increasing the number of vehicles reduces the rate of crashes. The results also suggest that the probability of crashes starts declining beyond the vertex of a parabola, where the parabola crosses its axis of symmetry. This model implies the probability for a crash to occur begins to decline by increasing the AADT beyond the 13,861 vertex. Congestion causes roads to operate at a fundamentally different level compared to pre-congestion traffic counts. The SIC effect may be a result of the characteristics of highly congested roads. The same trend exists for the bicyclist traffic regressor. An increase in traffic of bicyclists beyond the 1,314 vertex point decelerates the probability for a crash to occur. Bicyclists experience the SIN effect after such flows have been reached.

Looking at the second-part of the model, like the first-part, the downward parabola form of the *Vehicle Traffic* variable shows that the number of crashes has a diminishing return to scale considering the annual average daily traffic as an input. This is also true for the bicyclist traffic. Extracting the vertex points of both the exogenous variables, we found that the number of crashes starts decreasing when vehicle traffic and bicyclist traffic per intersection exceed 29,568 and 1,532, respectively.

Sensitivity Analysis

To quantify the association between number of crashes and interest variables, we calculated the elasticity of each independent variable. In line with our hypotheses, both the *Vehicle Traffic* and *Bicyclist Traffic* variables have an inelastic effect. The elasticity calculation indicates:

- **Vehicle Traffic**: A 1% increase in the annual average daily motor vehicle traffic increases the probability of crashes by 0.14% and the number of crashes, given there is a crash, by 0.80%.

- **Bicyclist Traffic**: A 1% increase in the annual average daily bicyclist traffic increases the probability of crashes by 0.09% and the number of crashes, given there is a crash, by 0.50%.

The two-part model demonstrates that vehicle and bicyclist traffic differ in their contribution to the probability and later the number of crashes. It appears that once an intersection has been predicted to have a crash by the Probit model, the effect of motor vehicle traffic on the number of crashes is greater than the effect of bicyclist traffic.

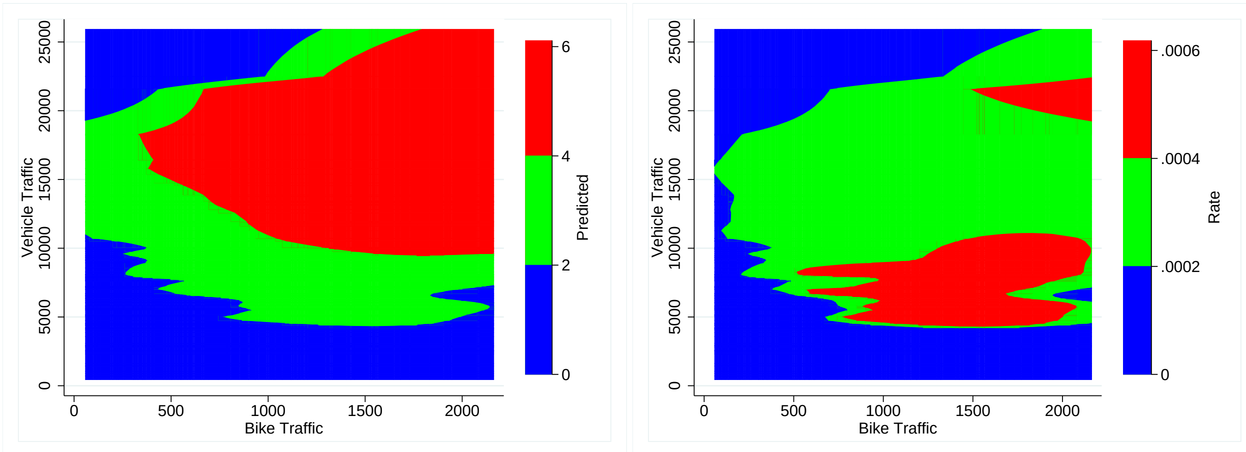
A key takeaway from the sensitivity analysis is that increasing the presence of motor vehicle traffic has a greater impact on the number of crashes than bicyclist traffic. It may be justified by the positive correlation between the bicyclist demand and bicyclist facility, which may result in more awareness of drivers.

Prediction Accuracy

Out of 106 test data, 55 intersections were observed to have between one and 11 crashes over the 14 year period. The first-part of the model predicted the 51.8% of the crashes accurately with the probability of 50%. The model is better than random with a Pseudo R^2 value of 0.03. A low R^2 value is acceptable in this case because bicyclist-auto crash occurrences are highly random events. To measure the prediction accuracy of the second-part of the model, we used the Mean Relative Percentage Error (MRPE) measurement as Equation 1. In this equation, C_i is the observed number of crashes at intersection i, $C_i^{'}$ is the predicted number of crashes at intersection i, and n stands for the number of observations.

$$\frac{1}{n}\sum_{i=1}^{n}\frac{\left|C_i^{'}-C_i\right|}{C_i}\times 100 \tag{1}$$

The MRPE results show that the second-part of the model predict the number of crashes with a 82.6% error on average. To graphically represent the prediction of crashes, we depict the the contour plots in Figure 4a and Figure 4b. These plots were generated by using the test dataset and applying the two-part model to predict the number of bicyclist-auto crashes across a range of bicyclist and vehicle traffic levels. These graphs are easy to read and can be used by engineers and planners alike to assess the vehicle and bicyclist risk associated with an intersection of recorded traffic flows.



(a) The predicted crashes

(b) The rate of number of crashes to traffic flow

FIGURE 4 : The contour plot

Intersections with one or two observed crashes tend to be overestimated in the model while intersections that have been observed to be more dangerous are underestimated.

SUMMARY AND CONCLUSIONS

The concept of safety in numbers in transportation planning reflects that there is a non-linear statistical correlation between the number of pedestrians and cyclists and the number of crashes. Studies used longitudinal and cross-sectional data at different level of aggregation to examine whether and to what extent the safety in numbers phenomenon is legitimate. The current study applies a two-part model of crashes on traffic data for 489 intersections in Minneapolis - St. Paul metropolitan area between 2000 and 2013.

We randomly divided the data into two sets to not only calibrate the model for number of crashes against the annual average daily vehicle traffic and the daily bicyclist traffic, but also to measure the accuracy of the model. To understand the association function between the number of crashes and both the annual average daily vehicle traffic and the daily bicyclist traffic, we embedded the quadratic functional form of our interest variables in the model. This enables us to shed light on the accuracy of the safety in numbers phenomenon and to quantify the safety returns to scale.

From the modeling side, we found that both the AADT and DBT has a diminishing return to scale. This accentuates the positive role of safety in numbers. Increasing the number of vehicles and cyclists decelerates not only the probability of crashes, but the number of crashes as well. However, their impacts are unequal. Measuring the elasticity of the variables, it is found that a 1% increase in the annual average daily motor vehicle traffic increases the probability of crashes by 0.14% and the number of crashes by 0.80%. However, a 1% increase in the average annual daily bicyclist traffic increases the probability of crashes by 0.09% and the number of crashes by 0.50%. We also found the saturation point of the safety in numbers for bicyclists is remarkably less than motor vehicles. Extracting the vertex point of the parabola functions reveals that the number of crashes starts decreasing when vehicle traffic and bicyclist traffic per intersection exceed 29,568 and 1,532, respectively.

As this study contemplated whether and to what extent the vehicle and bicyclist traffic affects the number of crashes, this provides insights for future research avenues. The following suggestions are made for further research:

- The use of additional road geometry features such as signalization and the number of approach lanes may improve the model, which results in getting a more accurate estimate of the safety in numbers effects.
- By accounting for variables that may influence vehicle and bicyclist traffic, we may explain a greater percentage of the variation in the number of crashes for a given intersection configuration and activity level.
- One caveat with the bicyclist-auto crash dataset is that bicyclists tend to report only more severe crashes. This means the crash records underreport the actual number of crashes that occur on a yearly basis which masks the true risk level at an intersection. The prevailing limitation to this and other bicyclist behavior studies, is the lack of consistent bicyclist TMC data. New sensors should make continuous counting of bicycle and vehicle traffic more standard.

REFERENCES

[1] Tami Toroyan, Margie M Peden, and Kacem Iaych. WHO launches second global status report on road safety. *Injury prevention : journal of the International Society for Child and Adolescent Injury Prevention*, 19(2):150, 2013.

[2] Walker Angell. Bicycling: Relatively Safe, 2014.

[3] Søren Underlien Jensen. Safety effects of blue cycle crossings: A before-after study. *Accident Analysis and Prevention*, 40(2):742–750, 2008.

[4] RJ Smeed. Some statistical aspects of road safety research. *J R Stat Soc A*, pages 1–34, 1949.

[5] Lars Leden. Pedestrian risk decrease with pedestrian flow. A case study based on data from signalized intersections in Hamilton, Ontario. *Accident Analysis and Prevention*, 34(4):457–464, 2002.

[6] L. Ekman. On the treatment of flow in traffic safety analysis-a non-parametric approach applied on vulnerable road users. *Lunds Tekniska Hogskola Bulletin*, 136, 1996.

[7] P L Jacobsen. Safety in numbers: more walkers and bicyclists, safer walking and bicycling. *Injury Prevention*, 9(3):205–209, 2003.

[8] JW Knott. Road traffic accidents in New South Wales, 1881-1991. *Australian Economic History Review*, 34:80–116, 1994.

[9] Rune Elvik. The non-linearity of risk and the promotion of environmentally sustainable transport. *Accident Analysis and Prevention*, 41(4):849–855, 2009.

[10] Rajiv Bhatia and Megan Wier. "Safety in Numbers" re-examined: Can we make valid or practical inferences from available evidence? *Accident Analysis and Prevention*, 43(1):235–240, 2011.

[11] Wesley E Marshall and Norman W Garrick. Evidence on why bike-friendly cities are safer for all road users. *Environmental Practice*, 13(1):16–27, 2011.

[12] Peter Lyndon Jacobsen, David R Ragland, and Charles Komanoff. Safety in numbers for walkers and bicyclists: exploring the mechanisms. *Injury prevention*, 21(4):217–220, 2015.

[13] Rune Elvik and Torkel Bjørnskau. Safety-in-numbers: a systematic review and meta-analysis of evidence. *Safety science*, 92:274–282, 2017.

[14] Dorothy L Robinson. Safety in numbers in australia: more walkers and bicyclists, safer walking and bicycling. *Health promotion journal of Australia*, 16(1):47–51, 2005.

[15] Judy Geyer, Noah Raford, Trinh Pham, and David Ragland. Safety in numbers: data from oakland, california. *Transportation Research Record: Journal of the Transportation Research Board*, (1982):150–154, 2006.

[16] Shane Turner, Graham Wood, Tim Hughes, and Rohit Singh. Safety performance functions for bicycle crashes in new zealand and australia. *Transportation Research Record: Journal of the Transportation Research Board*, (2236):66–73, 2011.

[17] Sandar Tin Tin, Alistair Woodward, Simon Thornley, and Shanthi Ameratunga. Regional variations in pedal cyclist injuries in new zealand: safety in numbers or risk in scarcity? *Australian and New Zealand journal of public health*, 35(4):357–363, 2011.

[18] Marilyn Johnson, Jennie Oxley, Stuart Newstead, and Judith Charlton. Safety in numbers? investigating australian driver behaviour, knowledge and attitudes towards cyclists. *Accident Analysis & Prevention*, 70:148–154, 2014.

[19] Jason Thompson, Giovanni Savino, and Mark Stevenson. Reconsidering the safety in numbers effect for vulnerable road users: an application of agent-based modeling. *Traffic injury prevention*, 16(2):147–153, 2015.

[20] K Todd. Pedestrian regulations in the United States: a critical review. *Transportation Quarterly*, 46:541–559, 1992.

[21] Robert J Schneider, Lindsay S Arnold, and David R Ragland. A Pilot Model for Estimating Pedestrian Intersection Crossing Volumes. 2010.

[22] Michael Iacono, Kevin J Krizek, and Ahmed El-Geneidy. Measuring Non-motorized Accessibility: Issues, Alternatives, and Execution. *Journal of Transport Geography*, 18(1):133–140, 2010.

[23] M. McDaniel, S., Lowry, M. B., Dixon. Using Origin-Destination Centrality to Estimate Directional Bicycle Volumes. page 15.

[24] Michael Lowry. Spatial interpolation of traffic counts based on origin–destination centrality. *Journal of Transport Geography*, 36:98–105, apr 2014.

[25] Luis F. Miranda-Moreno, Patrick Morency, and Ahmed M. El-Geneidy. The link between built environment, pedestrian activity and pedestrian-vehicle collision occurrence at signalized intersections. *Accident Analysis and Prevention*, 43(5):1624–1634, 2011.

[26] Maria Kamargianni and Amalia Polydoropoulou. Generation's Y Travel Behavior and Perceptions Towards Walkability Constraints. In *Transportation Research Board 93rd Annual Meeting*, volume 2430, pages 59–71, 2014.

[27] Steve Hankey and Greg Lindsey. Facility-demand models of peak period pedestrian and bicycle traffic: Comparison of fully specified and reduced-form models. *Transportation Research Record: Journal of the Transportation Research Board*, (2586):48–58, 2016.

[28] Andrew Owen, David Levinson, and Brendan Murphy. Access across america: Walking 2014. 2015.

[29] Jueyu Wang, Greg Lindsey, and Steve Hankey. Exposure to risk and the built environment, an empirical study of bicycle crashes in minneapolis. Technical report, 2017.

[30] Brendan Murphy, David M Levinson, and Andrew Owen. Evaluating the safety in numbers effect for pedestrians at urban intersections. *Accident Analysis & Prevention*, 106:181–190, 2017.

[31] Josh Frank Roll. *Bicycle Traffic Count Factoring: An Examination of National, State and Locally Derived Daily Extrapolation Factors*. PhD thesis, 2013.

[32] William H. Dow and Edward C. Norton. Choosing Between and Interpreting the Heckit and Two-Part Models for Corner Solutions. *Health Services and Outcomes Research Methodology*, 4(1):5–18, 2003.