Title
Unequal Streets: Active Transportation Safety Disparities in the SCAG Region

Author
O'Brien, Riley

Publication Date
2018

DOI
doi:10.17610/T63S32

Unequal Streets, Unequal Stations

Active Transportation Safety Disparities in the SCAG Region

PROJECT LEAD	Riley O'Brien
FACULTY SUPERVISOR	Professor Chris Tilly
CLIENT NAME	Tom Vo, Southern California Association of Governments

UCLA Institute of Transportation Studies

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
2018-54	N/A	N/A

4. Title and Subtitle Unequal Streets: Active Transportation Safety Disparities in the SCAG Region	**5. Report Date** 2018 **6. Performing Organization Code** UCLA-ITS
7. Author(s) Riley O'Brien	**8. Performing Organization Report No.** LAS1806
9. Performing Organization Name and Address Institute of Transportation Studies, UCLA 3320 Public Affairs Building Los Angeles, CA 90095-1656	**10. Work Unit No.** N/A **11. Contract or Grant No.** 2018-54
12. Sponsoring Agency Name and Address The University of California Institute of Transportation Studies www.ucits.org	**13. Type of Report and Period Covered** **14. Sponsoring Agency Code** UC ITS

15. Supplementary Notes

DOI: doi:10.17610/T63S32

16. Abstract

This study uses multivariate regression to determine correlates of pedestrian- and bicycle-involved crashes based on 14 demographic and built environment factors. At the census tract level, higher poverty rate and larger Hispanic/Latino population are statistically significant crash predictors of both pedestrian and bicycle crashes, even when controlling for job density, commute mode, and other variables related to pedestrian and bicycle exposure. Within ¼ mile of Metro stations and within ½ mile of Metrolink stations, the poverty rate is a significant predictor of at least one crash type. Within 1/4 mile of Metro stations, vehicle-miles traveled is a significant crash predictor, although not within ½ mile of Metrolink stations. These findings support increased investment in active transportation, including targeted funds for low-income community-led projects and projects near bus stops.

17. Key Words	18. Distribution Statement
vision zero equity pedestrian planning complete streets bicycle planning	No restrictions.

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	77	N/A

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

Unequal Streets, Unequal Stations

Active Transportation Safety Disparities in the SCAG Region

A comprehensive project submitted in partial satisfaction of the requirements for the degree Master of Urban and Regional Planning.

May 24, 2018

Author Riley O'Brien

Client Advisor Tom Vo, *Southern California Association of Governments*

Faculty Advisor Chris Tilly, *Professor of Urban Planning, UCLA*

Disclaimer

This report was prepared in partial fulfillment of the requirements for the Master in Urban and Regional Planning degree in the Department of Urban Planning at the University of California, Los Angeles. It was prepared at the direction of the Department and of the Southern California Association of Governments as a planning client. The views expressed herein are those of the authors and not necessarily those of the Department, the UCLA Luskin School of Public Affairs, UCLA as a whole, or the client.

Executive Summary

This study analyzes socioeconomic and built environment predictors of pedestrian- and bicycle-involved crashes in the six-county jurisdiction of the Southern California Association of Governments (SCAG). As the nation's largest metropolitan planning organization, with a per-capita pedestrian fatality rate higher than the state and national averages, SCAG has a unique responsibility to improve active transportation conditions. SCAG plays a direct role in allocating a large portion of statewide Active Transportation Program (ATP) funding, which has the potential to prevent injuries and deaths, reduce automobile use, and promote environmental justice more broadly. By providing clear evidence that high-poverty communities experience a disproportionate share of crashes, this report demonstrates the importance of the ATP and provides justification for its expansion.

I developed six linear regression models to identify predictors of pedestrian- and bicycle-involved crashes at three geographic scales. I considered 14 possible predictors, including built environment factors such as schools and commercial land use, transportation variables such as transit stops and vehicle miles traveled, density variables such as the number of people and jobs, and socioeconomic variables such as poverty rate and Hispanic/Latino population share. Using geographic information systems (GIS) software, I aggregated pedestrian- and bicycle-involved crashes and predictor variables to all census tracts in the SCAG region, as well as ¼ mile buffers around the region's Metrolink commuter rail stations and Los Angeles County's Metro rail and busway stations. Using statistical software, I conducted each regression to determine the relationship between each crash type and each predictor variable while controlling for the remaining predictor variables.

My regression results suggest that more vulnerable communities have less safe conditions for walking and biking, especially at the census tract level. The tract-level models account for 57% of the variation in pedestrian crashes and

49% of the variation in bicycle crashes, with all predictor variables statistically significant at a 95% confidence level. A higher poverty rate and Hispanic/Latino population share predict more bicycle and pedestrian crashes per tract, and a higher Black/African-American population share also predicts more pedestrian crashes per tract. The number of major transit stops per tract is the top predictor of pedestrian crashes and third-strongest predictor of bicycle crashes. These trends are less consistent in the station-level regression results, yet poverty is still one of the strongest predictors of both crash types near Metrolink stations and bicycle crashes near Metro stations. At the Metro station level, vehicle-miles traveled is also one of the strongest predictors of both crash types.

Based on my findings, I recommend expanding the required 25% allocation of ATP funds to disadvantaged communities as defined statewide, as well as prioritizing projects that would enhance connectivity, increase visibility, and reduce automobile speeds near major transit stops. I also recommend collecting additional data on bicycle and pedestrian activity, either by purchasing existing data or using automated counters in high-crash locations. For future research, I suggest using more advanced statistical techniques such as cluster analysis, incorporating bike lane and sidewalk data in the analysis, and conducting field audits and community workshops.

Table of Contents

Introduction

Traffic collisions are just one example of the negative externalities resulting from motorized transportation, along with noise, congestion, localized air pollution, and greenhouse gas emissions. Although some crashes involve only non-motorized modes, most pedestrian and bicycle crashes involve automobiles and other large motorized vehicles (Tuckel et al., 2014). The large speed and mass differential between motor vehicles and non-motorized modes make pedestrian-automobile and bicycle-automobile collisions particularly dangerous. For this reason, pedestrians, bicyclists, and other travelers not shielded by a heavy vehicle are "vulnerable road users", with children, the elderly, and people with disabilities particularly at risk (Organization for Economic Cooperation and Development, 1998). Racial and economic inequality exacerbate these vulnerabilities, as low-income people and people of color tend to have limited automobile access and therefore rely on other modes of travel including walking and biking (Blumenberg, 2017). Additionally, low-income people and people of color tend to have worse access to medical care (National Research Council, 2004), which makes survival and recovery from crashes more difficult. Therefore, unsafe conditions for active transportation worsen racial and economic inequality.

While reducing pedestrian and bicycle collisions should be a priority everywhere, reducing them in Southern California has unique importance. In 2016, California had the 10th most pedestrian fatalities per resident in the United States. The counties of Los Angeles, San Diego, and Orange had the highest number of pedestrian fatalities in the state (Retting, 2018). The Southern California Association of Governments (SCAG), the largest Metropolitan Planning Organization in the US, guides planning for Los Angeles and Orange counties, along with the counties of Riverside, San Bernardino, and Ventura. As Figure 1 shows, Los Angeles, Riverside, and San Bernardino counties had higher pedestrian fatality rates than California overall, and all SCAG counties except Ventura had higher pedestrian fatality rates than the national average.

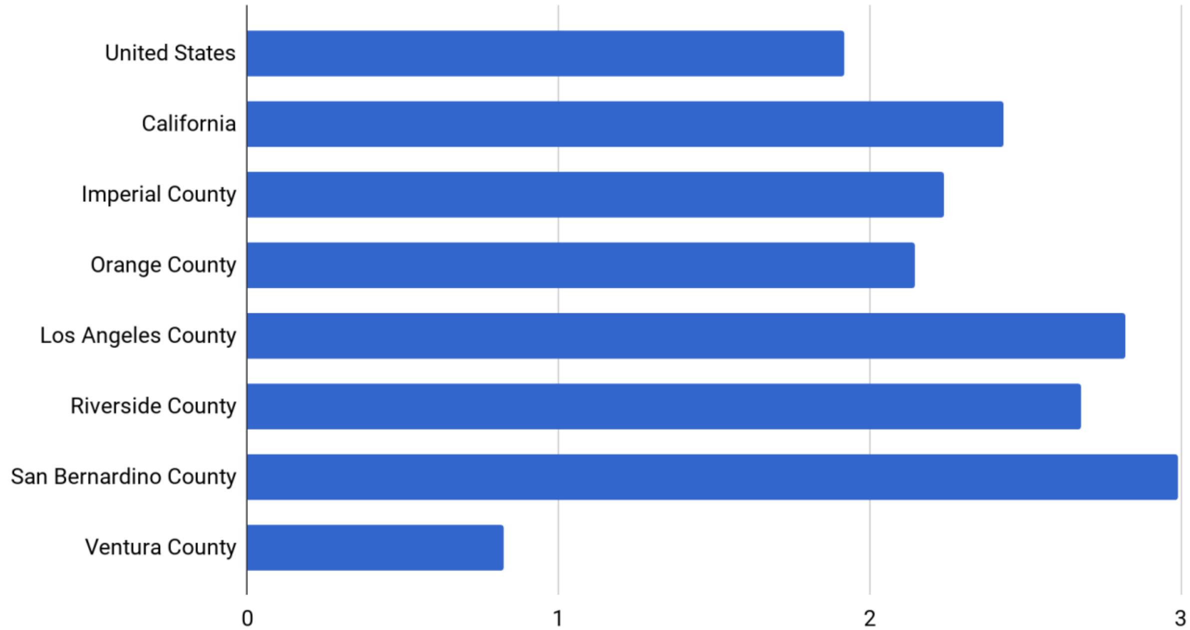

Figure 1: Pedestrian fatalities per 100,000 population in 2016 (Governors Highway Safety Association, 2018, California Highway Patrol, 2016, US Census Bureau, 2016)

On the surface, the relatively low per-capita fatality rates of Orange and Ventura counties appear to be success stories, but they may simply reflect lower rates of walking and higher rates of driving, which California transportation policy is increasingly focused on changing. In accordance with federal and state requirements, SCAG prepares a Regional Transportation Plan and Sustainable Communities Strategy (RTP/SCS) every four years outlining a path to compliance with state and federal environmental goals. SCAG's 2016 RTP/SCS allocates $12.9 billion to improve active transportation, anticipating a 28% increase in walking and 71% increase in biking throughout the region by 2040

(Southern California Association of Governments, 2016a). Achieving these lofty goals will require officials to prioritize the most efficient and equitable active transportation projects and build the necessary political coalitions to complete these projects. This study supports both objectives by identifying high-collision areas, ranking which factors predict crashes, and demonstrating that these areas tend to be low-income communities and communities of color.

Literature Review

Using a variety of methods, researchers in public health, transportation planning, and other fields have investigated where, when, and how active transportation collisions occur. These studies often use collision data gathered by law enforcement agencies, which generally omit sensitive information such as a victim's race, ethnicity, income, and residence. However, by linking collision location to various demographic and built environment characteristics, these studies have revealed higher pedestrian and bicycle collision rates in low income communities and communities of color, even when controlling for factors connected to higher pedestrian and bicycle activity levels. Although residents in these areas are not necessarily involved in these collisions, living in an area that is more dangerous for pedestrians and cyclists results in a lower quality of life. This represents a clear example of an environmental injustice, as lower income people and people of color are disproportionately harmed by unsafe conditions for walking and biking.

Spatial and Socioeconomic Variation in Crash Frequency

To determine which areas are most in need of safety improvements, numerous researchers have examined the relationship between pedestrian and/or bicycle collisions and characteristics of the surrounding area. Generally, these studies aggregate the number of bicycle and/or pedestrian-involved collisions to a defined spatial area and construct a multivariate regression model to determine which built environment and demographic factors most strongly predict collision frequency. Many of these studies use ordinary least squares (OLS) linear regression (Rivara and Barber, 1985, LaScala et al., 2000, Loukaitou-Sideris et al., 2007, Zhu and Lee, 2008, Weir et al., 2009, Sebert Kuhlmann et al. 2009), while several others use a negative binomial regression to account for the high number of census tracts with low numbers of collisions (Graham and Glaister, 2003, Kim et al., 2006, Chen et al., 2011, Pharr et al., 2013, Ukkusuri et al., 2012, Noland et al., 2013). Other regression methods include generalized linear models (GLM) (David and Rice, 1994, Wedegama et al., 2006), Poisson models (Cottrill and Thakuriah, 2010), and the M-Plus 6.11 model (Yu, 2014).

Ten of these studies focus only on pedestrian collisions, while four focus on both pedestrian and bicycle collisions and two focus only on collisions involving child pedestrians, with the time of observations ranging from one to seven years. Eleven use the United States census tract as the unit of analysis, while the remaining five studies use either neighborhoods, census block groups, British enumeration districts, British wards, or 0.1 square-mile grids as the unit of analysis. Thirteen studies analyzed collisions in the United States, two studies analyzed collisions in the United Kingdom, and one study analyzed collisions in Montreal, Canada.

Multiple studies revealed population, population density, employment density, traffic volume, land use, race, ethnicity, age, income, commute mode share, vehicle ownership, and street network design as significant predictors of the number of pedestrian and/or bicycle collisions in an area. A higher population often predicts a higher number of collisions (LaScala et al., 2000, Kim et al., 2006, Weir et al., 2009), whereas population density has a less clear relationship with collision frequency. Most studies examining population density have found a positive association between density and collision frequency (Graham and Glaister, 2003, Wedegama et al., 2006, Sebert Kuhlmann et al., 2006, Loukaitou-Sideris et al., 2007, Ukkusuri et al., 2012, Yu, 2014), although some have found the opposite (Pharr et al., 2013, Noland et al., 2013).This suggests a non-linear relationship between population density and collision frequency, in which moderate densities result in more pedestrian and bicycle activity and in turn more collisions, but higher densities are associated with slower automobile speeds and more safety measures. The latter trend is consistent with research demonstrating that there is “safety in numbers” when walking or biking, meaning more pedestrians and cyclists enhance driver awareness and safety (Jacobsen, 2003). Additionally, higher densities may predict fewer collisions because drivers in high-density areas drive slower and less often than drivers in low-density suburbs (Ewing and Kreutzer, 2006). This also underscores the difficulty of analyzing collision predictors when activity levels are not known.

Other characteristics that imply more pedestrian and bicycle activity have a more consistent positive relationship with collision frequency based on previous research. Higher employment density, measured as the number of jobs in a given area, frequently predicts higher collision rates (Graham and Glaister, 2003, Loukaitou-Sideris et al., 2007, Weir et al., 2009, Noland et al., 2013). The percentage of land use classified as commercial, representing locations where customers would be expected to walk and bike in addition to workers, also predicts higher collision frequency (David and Rice, 1994, Kim et al., 2006, Wedegama et al., 2006, Loukaitou-Sideris et al., 2007, Weir et al., 2009, Ukkusuri et al., 2012). Other land use characteristics are predictors of higher pedestrian and bicycle collision rates, including the number of transit stops (Ukkusuri et al., 2012, Pharr et al., 2013, Yu 2014) and the number of schools (Cottrill and Thakuriah, 2010, Pharr et al., 2013). Road network characteristics also play an important role, as a higher number of intersections predicts more collisions (Graham and Glaister, 2003, Wedegama et al., 2006, Yu, 2014), as does a higher proportion of streets that are one-way (David and Rice, 1994, Pharr et al., 2013). Automobile traffic volume, generally measured as average annual daily traffic (AADT), predicts more collisions according to several studies (LaScala et al., 2000, Loukaitou-Sideris et al., 2007, Weir et al., 2009, Cottrill and Thakuriah, 2010). Transportation habits also play a role, with a higher walking commute mode share predicting more collisions (Sebert Kuhlmann et al., 2009, Cottrill and Thakuriah, 2010) and a higher vehicle ownership rate predicting fewer collisions (Cottrill and Thakuriah, 2010, Noland et al., 2013).

Although many built environment characteristics influence spatial variation in pedestrian and bicycle collisions, they have not overshadowed the role of socioeconomic characteristics in the literature. The most commonly reported social predictor of pedestrian and bicycle collisions is the poverty rate, which tends to be higher in areas with more collisions (Rivara and Barber, 1985, Zhu and Lee, 2008, Weir et al., 2009, Yu, 2014). Similarly, a higher median household income predicts lower collision frequency (Rivara and Barber, 1985, Pharr et al., 2013). Age is also an important factor - a higher proportion of residents above age 65 predicts fewer collisions (Weir et al., 2009, Chen et al., 2011, Ukkusuri et al., 2012), while a higher proportion of children predicts fewer collisions in two studies (LaScala et al., 2000, Pharr et al., 2013) but more collisions in one study (Ukkusuri et al, 2012). Race and ethnicity also predict collisions, as a higher percentage of residents identifying as white predicts fewer collisions (Rivara and Barber, 1985, Yu, 2014), while a higher percentage of residents identifying as Hispanic or Latino predicts more collisions (Loukaitou-Sideris et al., 2007, Zhu and Lee, 2008, Pharr et al., 2013). The significance of demographic factors in predicting collisions shows that collision risk extends beyond physical factors and that disadvantaged populations are generally more vulnerable to collisions.

Possibilities and Limitations of Crash Reduction Measures

To address spatial variation in collisions, researchers have emphasized the role of travel patterns, land use, and infrastructure, and variations in these characteristics across space may explain inequity in collision risk. Litman (2011) estimates that a 1% reduction in vehicle-miles traveled should reduce the total number of crashes (including those only involving automobiles) by roughly 1.7%. Automobile speed is also a risk factor, as higher speeds result in more pedestrian injuries and fatalities (Ewing and Kreutzer, 2006). In addition to reducing posted speed limits, planners can lower vehicle speeds and reduce traffic collisions by changing the street environment. Narrowing lanes can reduce traffic collisions by forcing drivers to behave less aggressively, while reducing the number of lanes allows slower drivers to establish prevailing speeds (Ewing and Kreutzer, 2006). Infrastructure such as stop signs, roundabouts, and other traffic calming measures can also limit collisions by reducing vehicle speeds, with single-lane roundabouts representing the most effective upgrade (Retting et al., 2003).

Measures that separate pedestrians from automobiles are also important, with sidewalks, pedestrian-only signal phases, and pedestrian refuge islands most effective in reducing collisions (Retting et al., 2003). Researchers have also demonstrated that bicycle facilities such as dedicated lanes and physically-separated “cycletracks” also reduce bicycle collisions substantially (Reynolds et al. 2009). Better lighting and signage also reduces accidents by increasing pedestrian and cyclist visibility (Retting et al. 2003, Reynolds et al. 2009). Redesigning streets and installing infrastructure is expensive, and studies have shown lower income neighborhoods often have less complete infrastructure than higher income neighborhoods (Lowe, 2016).

Although investing in sidewalks, bike lanes, and other infrastructure in communities could prevent injuries and deaths, some researchers have been critical of the implementation of traffic safety infrastructure. Li and Joh (2017) find that investments in both mass transit and bicycle infrastructure contribute to rising property values, which they argue represents a benefit for cities investing in enhanced infrastructure. Yet McClintock (2017) calls the unintended consequences of similar investments in sustainability "ecogentrification", describing bike lanes, urban gardens, and similar public amenities as "cultural capital that a sustainable city's growth coalition in turn valorizes as symbolic sustainability capital used to extract rent and burnish the city's brand at larger scales." Conversely, Flanagan et al. (2016) raises the concern that rising incomes and property values cause, rather than result from, bicycle infrastructure investment. Comparing infrastructure and demographics across Chicago and Portland, they find that non-gentrifying neighborhoods receive less bicycle infrastructure than gentrifying neighborhoods with similar characteristics (Flanagan et al. 2016).

Other researchers have described similar dynamics at a local scale. Focusing specifically on a bikeway project in a gentrifying neighborhood in Portland, Lubitow and Miller (2013) demonstrate that many transportation officials view infrastructure in value-neutral terms, and that this emphasis on expertise encourages top-down planning and fosters resentment among some residents. While some research has offered further community engagement as a solution (Lubitow et al. 2016), other scholars have shown that the disconnect between active transportation advocates and their supposed beneficiaries reflects a much deeper problem in transportation advocacy of ignorance and perpetuation of racism (Lugo 2016). Additionally, bicycle advocates in the San Francisco Bay Area argued in the 1990s that infrastructure would promote economic growth, which Stehlin (2015) claims foreshadowed gentrification in the region. Together, these studies show that improving equity in active transportation requires more than infrastructure. Active transportation planners must ensure robust community participation in any project and must support measures to curb displacement near new infrastructure.

Research Design

Project Overview

I used a quantitative approach to investigate active transportation safety disparities in the SCAG region. Using Esri's ArcGIS software, I assembled collision, built environment, and socioeconomic data for the SCAG region's 3,913 census tracts, 108 Los Angeles County Metro rail and busway station areas, and 54 Metrolink commuter rail station areas. To determine how strongly these built environment and socioeconomic factors predict pedestrian and bicycle-involved crashes at each geographic scale (census tract, Metro station area, and Metrolink station area), I created six ordinary least squares (OLS) linear regression models. This also allowed me to control for factors that predict pedestrian and bicycle *activity* while comparing the relationship between crash frequency, poverty, race, and ethnicity. I focus on census tracts as well as station areas to determine if there are specific risk factors near rail stations, which are expanding in number across the region.

Dependent Variables

This study analyzes six dependent variables, corresponding to each of the six regression models - pedestrian-involved crashes per census tract, bicycle-involved crashes per census tract, pedestrian-involved crashes within ¼ mile of a Metro rail or busway station, bicycle-involved crashes within ¼ mile of a Metro station, pedestrian-involved crashes within ½ mile of a Metrolink commuter rail station, and bicycle-involved crashes within ½ mile of a Metrolink station. These point data were obtained from the California Highway Patrol (CHP)-maintained Statewide Integrated Traffic Records System (SWITRS) and geocoded UC Berkeley's Transportation Injury Mapping System (TIMS) for the years 2005 to 2014. All pedestrian or bicycle-involved crashes in the SCAG region during this period were included in the census tract models. All six dependent variables had a unimodal distribution and a positive skew, which I addressed by performing a log transformation of each dependent variable before conducting the regressions. As many tracts and station areas experienced zero collisions during the study period, and the logarithm of zero is undefined, I took the natural log of the number of crashes plus one.

Although there were 3,956 census tracts in the SCAG region in 2012, I deemed 43 tracts unsuitable for the regression models, leaving a total of 3,913 tracts. These include 32 tracts with less than 100 residents, 7 tracts for which poverty rate could not be determined by the Census Bureau and 3 tracts with no residents classified as workers. I also removed the census tract containing Los Angeles City Hall, as the LEHD data mistakenly classified all employees of the City of Los Angeles as working at City Hall.

In the Metro and Metrolink station-level models, I restricted the data to the years 2013-2014, as no new stations were opened during those years. I removed all Metro Orange and Silver Line stops that were not located along the separated busway, including Orange Line stops in Warner Center and Silver Line stops in San Pedro and Downtown Los Angeles. I also removed the Civic Center / Grand Park station, as it contains Los Angeles City Hall and thus distorts employment data as mentioned in the previous paragraph. For Metrolink, I removed all stations without weekday service from the analysis, excluding those with service only on weekends or during special events.

Independent Variables

I examined 15 independent variables as described in Table 1 and included 14 of them in each tract-level model and 13 in each station-level model. I excluded the percent of workers commuting by bicycle from the three pedestrian crash models and I excluded the percentage of workers commuting by walking or public transit from the three bicycle crash models. At the station area level, population and population density are redundant as each observation is a circle, with a radius of ¼ mile (for Metro stations) or ½ mile (for Metrolink stations). Therefore, I excluded total population from the four station area models.

Table 1: Independent variables in analysis

Independent Variable	Source	Geography	Category
Number of intersections	SCAG 2012 RTP/SCS	Point	Exposure
Number of K-12 schools	California School Campus Database (2016)	Point	Exposure
Percent of land use classified as commercial	SCAG 2012 RTP/SCS	Polygon (Parcel)	Exposure
Number of major transit stops	SCAG 2012 RTP/SCS	Point	Exposure
Vehicle-miles traveled	Internal SCAG travel model (2016)	Polyline (Road Segment)	Exposure
Number of jobs	Longitudinal Employer - Household Dynamics (2012)	Point	Exposure
Total population*	American Community Survey 2012 5-Year Estimates	Polygon (Census Tract)	Exposure

Population density	American Community Survey 2012 5-Year Estimates	Polygon (Census Tract)	Exposure
Percent of households with child under 18	American Community Survey 2012 5-Year Estimates	Polygon (Census Tract)	Exposure
Percent of residents over 65	American Community Survey 2012 5-Year Estimates	Polygon (Census Tract)	Exposure
Percent of workers commuting by walking or public transit**	American Community Survey 2012 5-Year Estimates	Polygon (Census Tract)	Exposure
Percent of workers commuting by bicycle***	American Community Survey 2012 5-Year Estimates	Polygon (Census Tract)	Exposure
Percent of residents identifying as Black or African-American	American Community Survey 2012 5-Year Estimates	Polygon (Census Tract)	Environmental Justice
Percent of residents identifying as Hispanic or Latino	American Community Survey 2012 5-Year Estimates	Polygon (Census Tract)	Environmental Justice
Poverty rate	American Community Survey 2012 5-Year Estimates	Polygon (Census Tract)	Environmental Justice

*tract-level models only   **pedestrian crash models only   ***bicycle crash models only

Although my crash data spans 2005 to 2014, SCAG data is generally available in 4-year increments corresponding with the RTP/SCS cycle. For that reason, 2008 and 2012 are the only years in which data is available within the crash data timespan. Although 2008 is closer to the center of the crash data timespan, I chose to use 2012 data wherever possible, as 2008 represents an outlier due to the impact of the Great Recession, especially on the number of jobs. I used 2016 data for schools and vehicle-miles traveled as 2012 data for these variables was not readily available.

I used the ArcGIS spatial join feature to link these variables to each census tract and station-area buffer. This was straightforward for point data, but more complicated for other data types. For vehicle-miles traveled polyline data, I used the "Feature to Point" tool to convert each road segment into its midpoint before joining each midpoint to the corresponding geographic unit. For percent of land use classified as commercial, I filtered out non-commercial land uses from SCAG's 2012 existing land use parcel-level dataset. After intersecting these parcels with the census tracts or station buffers, I divided the total area of commercial land use by the total area of the tract or buffer to obtain the proportion of commercial land use.

As American Community Survey data exist at the census tract level, I allocated these data to station buffers proportionally according to the area of overlapping census tracts. To estimate the population within each station buffer, I multiplied the population of each overlapping census tract by the proportion of the tract inside the buffer and added the resulting populations. To estimate ratio data such as population density and commute mode, I multiplied each census tract's ratio by the proportion of the buffer covered by the census tract. These proportional allocation methods assume that the population within each census tract is uniform and thus may differ from the actual station-area characteristics.

Many of the independent variables are positively skewed, and as a result I took the natural logarithm of these data (plus one, as the logarithm of zero is undefined) before conducting the regression. At the census tract level, all variables are positively skewed except four socioeconomic characteristics: under 18 population share, over 65 population share, Hispanic/Latino population share, and poverty rate. At the Metro and Metrolink station levels, the same variables were positively skewed except the number of intersections. However, I only used log transformations for my non-percent variables to facilitate interpretation of the regression results.

Limitations

My project has several limitations, including issues with data availability, data accuracy, and GIS analysis. The "'Exposure" variables identified in Table 1 are rough proxies for pedestrian and bicycle activity, and more accurate data on pedestrians and bicycles would allow me to differentiate between activity levels and safety levels. In addition, the SWITRS data only includes injuries which were reported to the authorities. Therefore, injuries affecting people fearful of the police, such as undocumented immigrants, may be missing from the SWITRS data. As I mentioned previously, I used 2016 data on the vehicle-miles traveled and the number of schools in each unit of analysis, unlike other independent variables which use 2012 data. Although traffic patterns and school locations were likely similar four years earlier, using 2016 data may underestimate the influence of these two variables on crash frequency.

Findings and Analysis

Crashes by Census Tract

Dependent Variables

Table 2 shows descriptive statistics for pedestrian-involved and bicycle-involved collisions at the census tract level between 2005 and 2014. While more pedestrian collisions occurred than bicycle collisions during that period, their totals, medians, means, and standard deviations were similar. For both crash types, the mean number of crashes per census tract exceeded the median, suggesting a positively-skewed distribution.

Table 2: Descriptive statistics for crashes by census tract

Variable	Total	Minimum	Maximum	Median	Mean	Standard Deviation
Pedestrian-involved crashes	64,692	0	294	12	17	17
Bicycle-involved crashes	55,897	0	256	10	14	15

Figures 2 and 3 show the distribution of pedestrian-involved crashes and bicycle-involved crashes, respectively, at the census tract level. These graphs confirm the positive skew suggested by each distribution's mean and median, and also reveal unimodal distributions, as the majority of the 3,913 census tracts experienced between 1 and 20 collisions across the 10-year period.

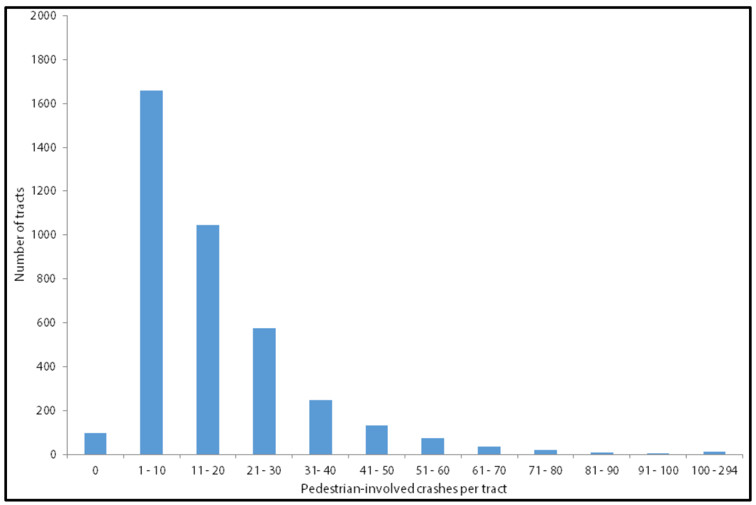

Figure 2: Histogram of pedestrian-involved crashes by census tract, 2005-2014

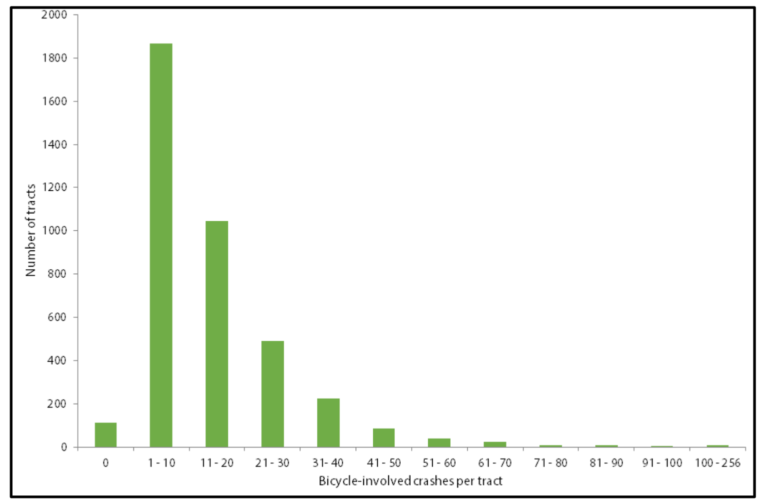

Figure 3: Histogram of bicycle-involved crashes by census tract, 2005-2014

Figures 4 and 5 show the distribution across the region of pedestrian- and bicycle-involved crashes, respectively. Census tracts are classified by quintile, with nearly all top 20% collision tracts in the most urban portions of the six SCAG counties for both crash types. Bicycle-involved crashes are somewhat more concentrated in urban areas than pedestrian-involved crashes.

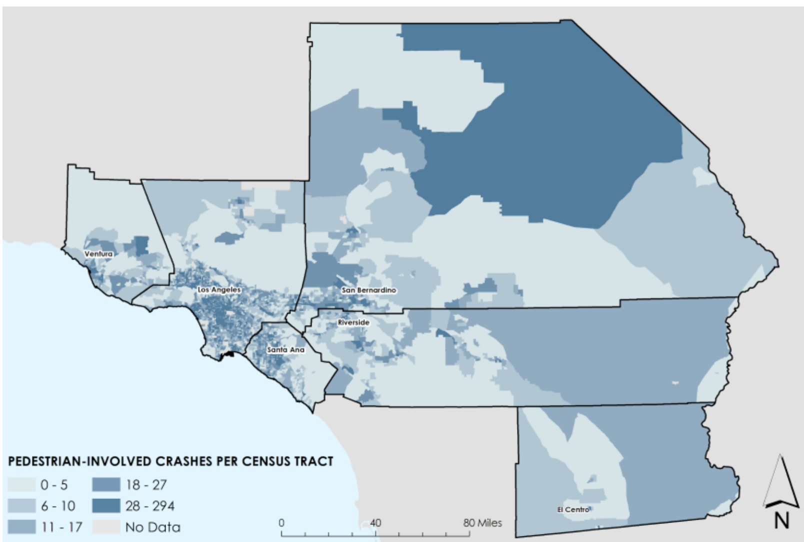

Figure 4: Map of pedestrian-involved crashes by census tract, 2005-2014

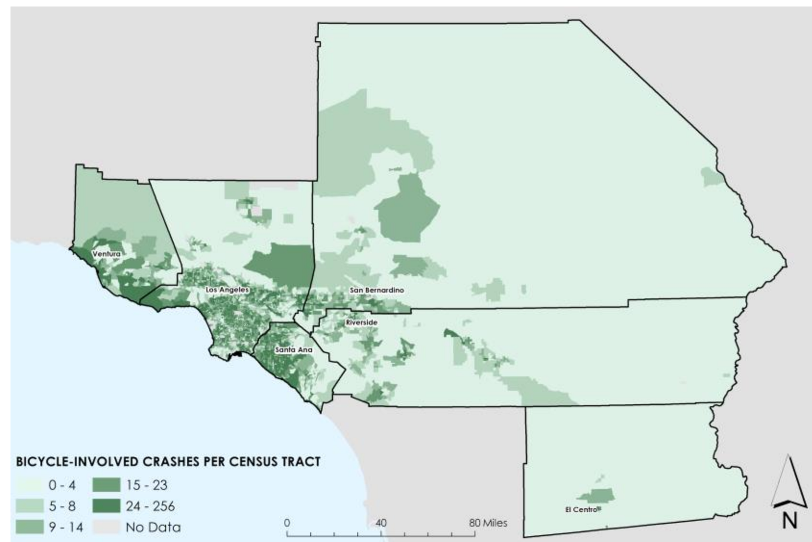

Figure 5: Map of bicycle-involved crashes by census tract, 2005-2014

Independent Variables

Table 3 includes descriptive statistics for the 12 independent variables. Most variables fluctuate greatly between tracts and appear to have a positive skew, as indicated by the standard deviation and mean exceeding the median.

Table 3: Descriptive statistics for independent variables by census tract

Variable	Total	Minimum	Maximum	Median	Mean	Standard Deviation
Intersections	461,792	0	2230	90	118	127
Schools	4,204	0	11	1	1	1
Commercial land use share	N/A	0%	72%	6%	8%	8%
Stops	157,608	0	1,308	23	40	64
Vehicle-miles traveled	175,000,000	0	677,676	31,597	44,628	49,502
Jobs	6,711,608	1	77,047	722	1,715	3,686
Population	18,000,000	101	22,123	4,403	4,613	1,875
Population per square mile	N/A	0.34	112,691	7,892	10,160	9,763
Under 18 population share	N/A	0	64%	18%	18%	6%
Over 65 population share	N/A	0	85%	10%	12%	8%
Black/African-American population share	N/A	0	93%	3%	6%	11%
Hispanic/Latino population share	N/A	0	100%	40%	44%	28%
Poverty rate	N/A	0	92%	13%	16%	12%
Combined walking and transit commute mode share	N/A	0	89%	5%	8%	10%
Bicycle commute mode share	N/A	0	20%	0%	1%	1%

Regionwide Crash Predictors

Pedestrian-Involved Crash Predictors

Table 4 displays the results of the linear regression model for pedestrian collisions at the census tract level. As the R^2 of 0.5673 indicates, approximately 57% of the variation in collision frequency across census tracts can be explained by the independent variables. All independent variables were statistically significant crash predictors at a 0.05 significance level. All independent variables predicted an increase in collision frequency except for the percent of residents below age 18 and above age 65, which is consistent with past research (LaScala et al., 2000, Weir et al., 2009, Chen et al., 2011, Ukkusuri et al., 2012, Pharr et al., 2013).

Whereas the unstandardized coefficients show how a 1-unit increase in the independent variable would change the number of crashes per tract, the standardized (beta) coefficients show how many standard deviations (SD) in crashes per tract a 1-SD increase in the independent variable would cause. Therefore, the standardized coefficients allow for comparison between independent variables. Based on this, the top three pedestrian collision predictors at the census tract level are the number of transit stops, the population per square mile, and the poverty rate. The predictive power of transit stops may simply result from high transit availability in areas with more walking, but it also suggests that motorists are colliding with transit riders on their way to the stop. Population density had a stronger impact on crash frequency than total population, which likely reflects conventional wisdom that denser areas have more pedestrian activity. The strong relationship between poverty and pedestrian crash frequency underscores that low income communities are disproportionately harmed by unsafe traffic conditions. This relationship's persistence despite so many control variables implies high-poverty communities are more likely to have higher rates of pedestrian activity for non-work trips, and/or unsafe behavior and infrastructure is more prevalent in high-poverty communities.

Table 4: Regression results for pedestrian-involved crashes by census tract

Independent Variable	Unstandardized (b) Coefficient	Standard Error	t	P > |t|	Beta
Number of intersections (natural log)	0.0700	0.0257	2.73	0.006	0.0599
Number of schools (natural log)	0.1066	0.0207	5.14	0.000	0.0601
Commercial land use share	0.0122	0.0016	7.54	0.000	0.1074
Major transit stops (natural log)	0.1676	0.0085	19.67	0.000	0.2637
Vehicle-miles traveled (natural log)	0.0657	0.0080	8.17	0.000	0.1026
Jobs (natural log)	0.1260	0.0119	10.56	0.000	0.1641
Population (natural log)	0.1217	0.0323	3.76	0.000	0.0595
Population per square mile (natural log)	0.1603	0.0135	11.88	0.000	0.2400
Under 18 population share	-0.0196	0.0023	-8.34	0.000	-0.1258
Over 65 population share	-0.0033	0.0016	-2.05	0.040	-0.0273
Combined walking and transit commute mode share	0.0091	0.0015	6.24	0.000	0.0942
Black/African-American population share	0.0070	0.0010	7.36	0.000	0.0835
Hispanic/Latino population share	0.0045	0.0005	8.50	0.000	0.1361
Poverty rate	0.0139	0.0012	11.42	0.000	0.1787
Number of observations	3,913	**R-squared**	0.5673	**Adjusted R-squared**	0.5658

Bicycle-Involved Crash Predictors

Table 5 displays the results of the linear regression model for bicycle collisions at the census tract level. As the R^2 of 0.4853 indicates, approximately 49% of the variation in collision frequency across census tracts can be explained by the independent variables. All independent variables were statistically significant crash predictors at a 0.01 significance level. All independent variables predicted an increase in collision frequency except for the percent of residents below age 18, the percent of residents above age 65, the percent of Black or African-American residents, and total population. The negative association between bicycle crash frequency and young or old age is consistent with aforementioned findings on pedestrian-involved crashes, likely reflecting the tendency for young and old people to ride bikes less. According to the standardized coefficients, the top 3 crash predictors are population density, number of intersections, and number of transit stops. The predictive power of population density and the number of transit stops is consistent with pedestrian-involved collisions, but the number of intersections stands out as influencing bicycle crashes far more than pedestrian crashes. This may suggest that Southern California drivers check for pedestrians more often than cyclists when turning at intersections.

Table 5: Regression results for bicycle-involved crashes by census tract

Independent Variable	Unstandardized (b) Coefficient	Standard Error	t	P > |t|	Beta
Number of intersections (natural log)	0.2803	0.0275	10.19	0.000	0.2411
Number of schools (natural log)	0.0713	0.0225	3.17	0.002	0.0404
Commercial land use share	0.0096	0.0018	5.45	0.000	0.0846
Major transit stops (natural log)	0.1386	0.0092	15.01	0.000	0.2189
Vehicle-miles traveled (natural log)	0.0992	0.0087	11.35	0.000	0.1555
Jobs (natural log)	0.1522	0.0130	11.72	0.000	0.1991
Population (natural log)	-0.1396	0.0351	-3.98	0.000	-0.0685
Population per square mile (natural log)	0.2968	0.0148	20.11	0.000	0.4461
Under 18 population share	-0.0110	0.0025	-4.38	0.000	-0.0709
Over 65 population share	-0.0059	0.0017	-3.42	0.001	-0.0499
Bicycle commute mode share	0.0739	0.0084	8.81	0.000	0.1077
Black/African-American population share	-0.0040	0.0010	-3.82	0.000	-0.0475
Hispanic/Latino population share	0.0015	0.0006	2.56	0.010	0.0448
Poverty rate	0.0086	0.0012	7.20	0.000	0.1112
Number of observations	3,913	**R-squared**	0.4853	**Adjusted R-squared**	0.4835

Crashes by Metro Rail/Busway Station

Dependent Variables

Table 6 shows descriptive statistics for crashes within ¼ mile of Metro rail and busway stations from 2013 to 2014. As the median is less than the mean and standard deviation for both crash types, the distribution appears to be positively skewed.

Table 6: Descriptive statistics of crashes by Metro station, 2013-2014

Variable	Total	Minimum	Maximum	Median	Mean	Standard Deviation
Pedestrian-involved crashes	696	0	33	4	6	7
Bicycle-involved crashes	672	0	25	4	6	6

Figures 6 and 7 are histograms of pedestrian- and bicycle-involved crashes, respectively, by Metro station area. They each confirm the positive skew implied by the descriptive statistics and reveal a unimodal distribution in which most stations experienced between 1 and 10 crashes during the two-year period.

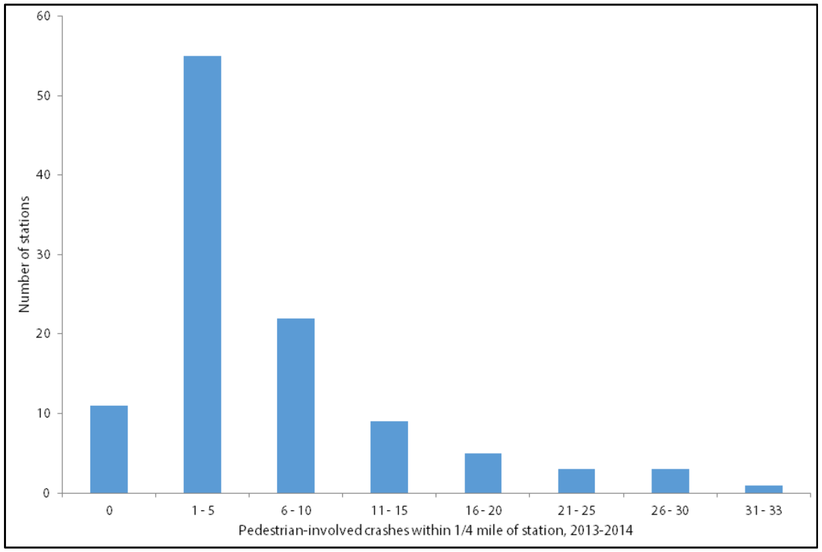

Figure 6: Histogram of pedestrian-involved crashes per Metro station, 2013-2014

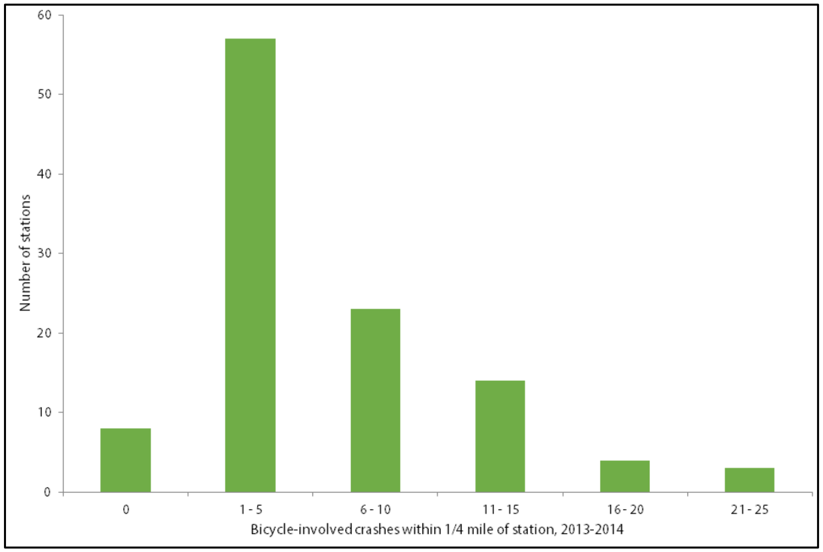

Figure 7: Histogram of bicycle-involved crashes per Metro station, 2013-2014

Figures 8 and 9 show the geographic distribution of pedestrian- and bicycle-involved crashes, respectively, by Metro station. The highest crash rates were in Central Los Angeles, with stations in Long Beach and South Los Angeles also exhibiting high crash frequency.

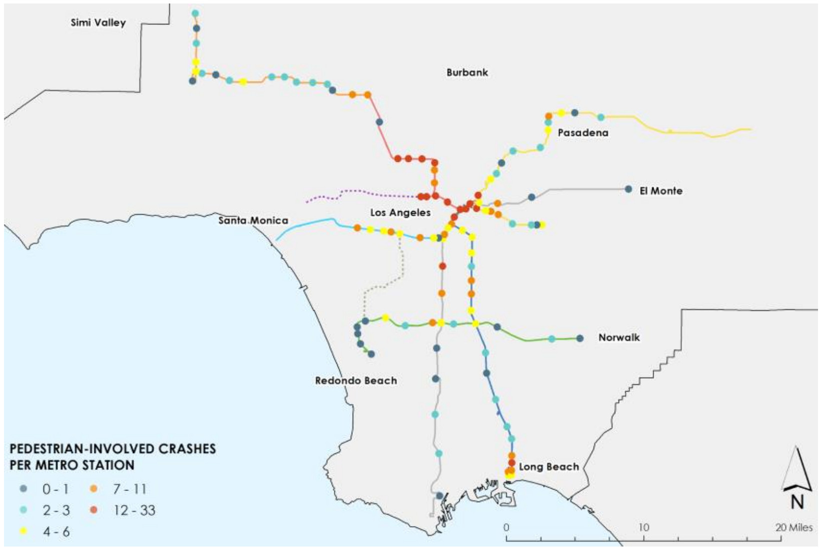

Figure 8: Map of pedestrian-involved crashes within ¼ mile of Metro stations, 2013-2014

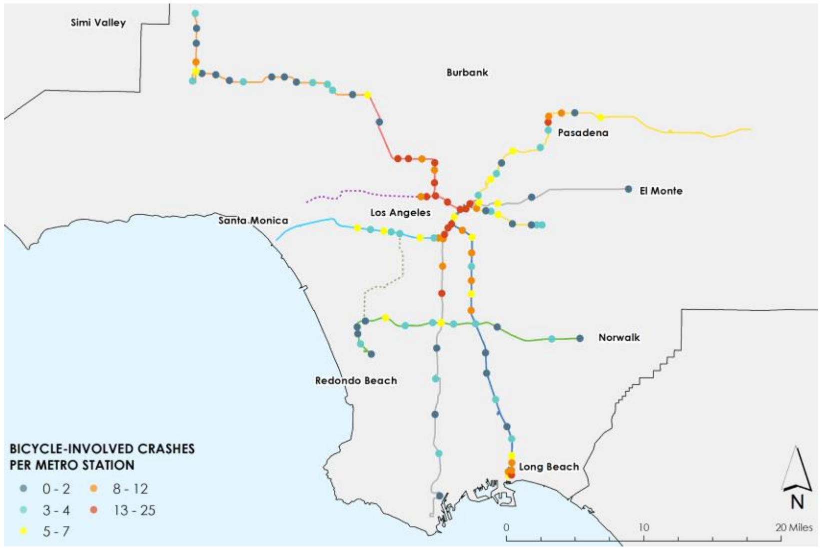

Figure 9: Map of bicycle-involved crashes within ¼ mile of Metro stations, 2013-2014

Independent Variables

Table 7 shows descriptive statistics for each independent variable within ¼ mile of a Metro station. Compared to the median census tract in the entire SCAG region, the median Metro station area has fewer intersections, fewer schools, a higher share of commercial land use, more transit stops, fewer vehicle-miles traveled, more jobs, more people per square mile, a lower percentage of people under 18 and over 65, and a higher percentage of people commuting by walking or transit, commuting by bicycle, identifying as Black or African-American, identifying as Hispanic or Latino, and living below the poverty line.

Table 7: Descriptive statistics of independent variables by Metro station area

Variable	Total	Minimum	Maximum	Median	Mean	Standard Deviation
Intersections	4,578	9	96	40	42	18
Schools	60	0	8	0	1	1
Commercial land use share	N/A	0%	82%	16%	20%	17%
Stops	13,522	4	1268	72	124	183
Vehicle-miles traveled	3,238,795	4381	80,359	23,738	29,714	17,263
Jobs	485,183	0	130,597	1,204	4,451	13,470
Population per square mile	N/A	0.00	50,156	11,549	13,150	8,651
Under 18 population share	N/A	0%	32%	15%	15%	7%
Over 65 population share	N/A	0%	29%	9%	10%	5%
Combined walking and transit commute mode share	N/A	0%	71%	15%	18%	13%
Bicycle commute mode share	N/A	0%	8%	1%	1%	1%
Black/African-American population share	N/A	0%	50%	6%	11%	12%
Hispanic/Latino population share	N/A	0%	99%	44%	49%	26%
Poverty rate	N/A	0%	66%	24%	24%	14%

Metro Station Crash Predictors

Pedestrian-Involved Crash Predictors

Table 8 displays the results of the linear regression model for pedestrian collisions within ¼ mile of Metro rail and busway stations in 2013 and 2014. As the R^2 of 0.6324 indicates, approximately 63% of the variation in collision frequency across census tracts can be explained by this model. Unlike the tract-level model, most of the independent variables in this model are not statistically significant, even at a 0.1 significance level. Only vehicle-miles traveled and combined walking and transit commute mode share are significant at a 0.01 significance level. The number of jobs is a significant predictor at the 0.05 level, while the percent of residents below age 18 and the population per square mile are both significant at the 0.1 level.

According to the standardized coefficients, the most influential predictors are the combined walking and transit commute mode share, vehicle-miles traveled, and the under 18 population share. The influence of commute mode is unsurprising yet intriguing, as it suggests that pedestrian crash victims near Metro stations are more likely to be on their way to work than victims in the region overall. Vehicle-miles traveled is also a much stronger predictor at the station level than at the tract level, implying that station-level traffic calming and even diversion of traffic onto alternate routes may reduce crashes. Finally, the influence of the under 18 population is surprising, especially as it predicts more crashes at the station level despite predicting fewer crashes at the tract level. This may be related to the demographics or habits of the station areas - either more youth live near the highest crash stations in Central and South LA, youth are more likely to walk near Metro stations than in other areas, or young pedestrians near Metro stations are less safe due to lacking infrastructure or education. Finally, although its positive association has a slightly more than 16% chance of being a result of random error, a higher Black or African-American population share predicts more crashes near Metro stations, suggesting that station areas with high Black or African-American populations may have worse infrastructure compared to areas with other demographic characteristics.

Table 8: Regression results for pedestrian-involved crashes by Metro station area

Independent Variable	Unstandardized (b) Coefficient	Standard Error	t	P > |t|	Beta
Number of intersections	0.0024	0.0034	0.68	0.496	0.0489
Number of schools (natural log)	0.0620	0.1260	0.49	0.624	0.0347
Commercial land use share	0.0068	0.0046	1.48	0.142	0.1278
Major transit stops (natural log)	-0.0033	0.0758	-0.04	0.966	-0.0036
Vehicle-miles traveled (natural log)	0.4738	0.1259	3.76	0.000	0.3156
Jobs (natural log)	0.1093	0.0528	2.07	0.041	0.1990
Population per square mile (natural log)	0.0870	0.0513	1.70	0.093	0.1460
Under 18 population share	0.0326	0.0179	1.83	0.071	0.2612
Over 65 population share	-0.0007	0.0140	-0.05	0.963	-0.0038
Combined walking and transit commute mode share	0.0327	0.0091	3.59	0.001	0.4816
Black/African-American population share	0.0078	0.0056	1.39	0.167	0.1095
Hispanic/Latino population share	-0.0056	0.0048	-1.18	0.242	-0.1616

Poverty rate	-0.0084	0.0079	-1.06	0.293	-0.1293
Number of observations	108	**R-squared**	0.6324	**Adjusted R-squared**	0.5816

Bicycle-Involved Crash Predictors

Table 9 displays the results of the linear regression model for bike crashes near Metro stations. As the R^2 of 0.6217 indicates, approximately 62% of the variation in collision frequency across census tracts can be explained by the independent variables. Similar to the Metro station level pedestrian crash model, less than half of the independent variables were statistically significant at the 0.1 level. Vehicle-miles traveled was significant at the 0.01 level, while population density, commercial land use share, and poverty rate were significant at the 0.05 level. Additionally, the number of jobs was a statistically significant predictor at the 0.1 level. According to the standardized coefficients, the three most influential predictors were vehicle-miles traveled, population density, and the poverty rate. The impact of vehicle-miles traveled suggests that traffic calming or diverting as mentioned in the previous section would also help prevent bicycle crashes. The influence of the poverty rate also underscores the inequitable impact of bicycle safety in the region, and shows that this disparity exists even at the station level.

Table 9: Regression results for bicycle-involved crashes by Metro station area

Independent Variable	Unstandardized (b) Coefficient	Standard Error	t	P > |t|	Beta
Number of intersections	-0.0024	0.0031	-0.77	0.445	-0.0558
Number of schools (natural log)	0.1594	0.1174	1.36	0.178	0.0984
Commercial land use share	0.9107	0.4234	2.15	0.034	0.1875
Major transit stops (natural log)	0.1123	0.0715	1.57	0.119	0.1366
Vehicle-miles traveled (natural log)	0.3481	0.1134	3.07	0.003	0.2555
Jobs (natural log)	0.0883	0.0477	1.85	0.068	0.1772
Population per square mile (natural log)	0.1215	0.0464	2.62	0.010	0.2247
Under 18 population share	-0.0005	0.0161	-0.04	0.971	-0.0051
Over 65 population share	-0.0044	0.0129	-0.34	0.733	-0.0283
Bicycle commute mode share	0.0542	0.0464	2.62	0.010	0.2247
Black/African-American population share	0.0002	0.0052	0.04	0.969	0.0031
Hispanic/Latino population share	0.0002	0.0042	0.04	0.967	0.0056
Poverty rate	0.0135	0.0062	2.17	0.032	0.2295
Number of observations	108	**R-squared**	0.6217	**Adjusted R-squared**	0.5694

Crashes by Metrolink Station

Dependent Variables

Table 9 shows descriptive statistics for crashes within ½ mile of Metrolink stations in the SCAG region. As is true at the census tract and Metro station levels, the median is less than the mean and the standard deviation, implying a positively skewed distribution for both crash types.

Table 10: Descriptive statistics for crashes by Metrolink station

Variable	Total	Mini - mum	Maxi - mum	Median	Mean	Standard Deviation
Pedestrian-involved crashes	325	0	35	4	6	7
Bicycle-involved crashes	311	0	22	3	6	6

Figures 10 and 11 are histograms of pedestrian- and bicycle-involved crashes, respectively, near Metrolink stations. They confirm the positive skew of both distributions while revealing unimodal distributions in which most stations have had 0-10 of each collision type.

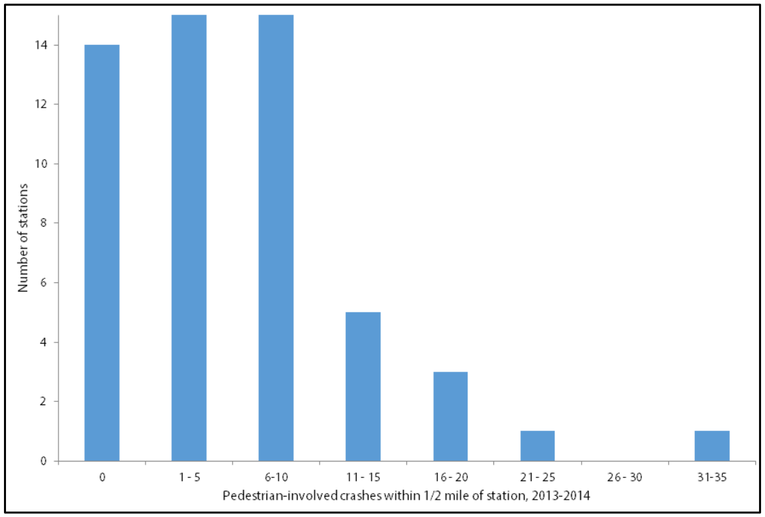

Figure 10: Histogram of pedestrian-involved crashes per Metrolink station, 2013-2014

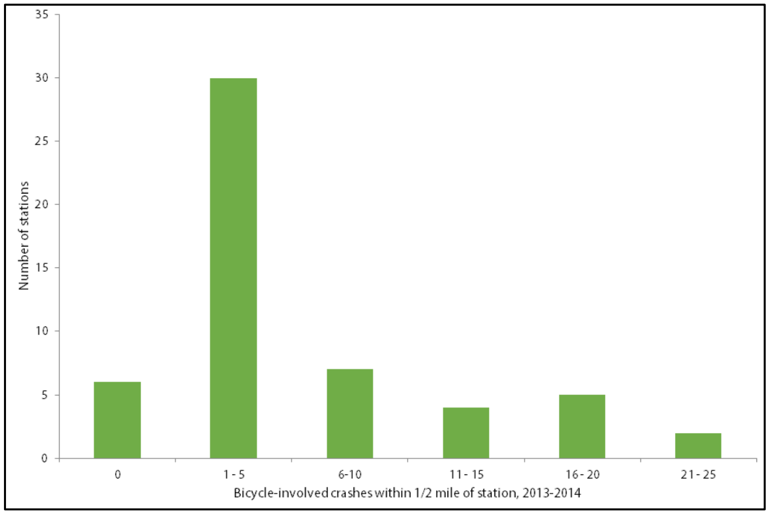

Figure 11: Histogram of bicycle-involved crashes per Metrolink station, 2013-2014

Figures 12 and 13 show the geographic distribution of pedestrian- and bicycle-involved crashes, respectively, by Metro station. The highest crash rates were in Central Los Angeles, with stations in Long Beach and South Los Angeles also exhibiting high crash frequency.

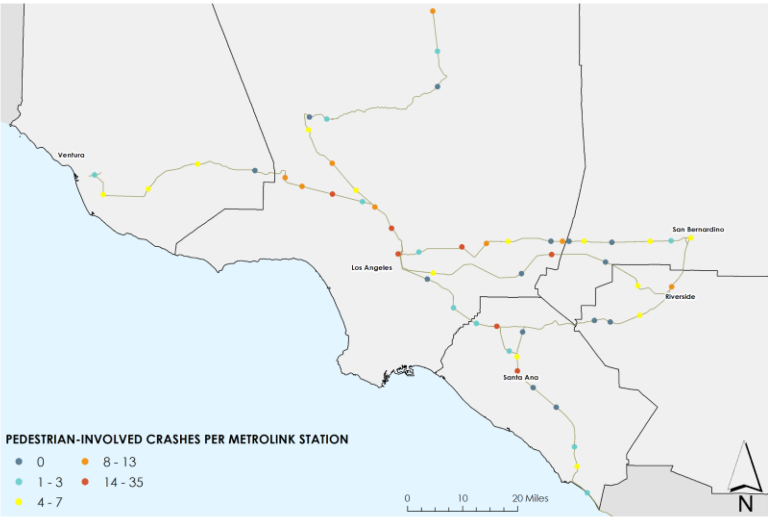

Figure 12: Map of pedestrian-involved crashes within ½ mile of Metrolink stations, 2013-2014

Figure 13: Map of bicycle-involved crashes within ½ mile of Metrolink stations, 2013-2014

Independent Variables

Table 10 shows descriptive statistics for each independent variable within ½ mile of a Metrolink station. Compared to the median census tract in the entire SCAG region, the median Metrolink station area has more intersections, fewer schools, a higher share of commercial land use, fewer transit stops, more vehicle-miles traveled, more jobs, fewer people per square mile, a similar percentage of people under 18, over 65, identifying as Black or African-American, and commuting by bicycle, and a higher percentage of people commuting by walking or transit, identifying as Hispanic or Latino, and living below the poverty line.

Table 11: Descriptive statistics of independent variables by Metrolink station area

Variable	Total	Minimum	Maximum	Median	Mean	Standard Deviation
Intersections	6,202	16	208	119	115	50
Schools	42	0	3	0	1	1
Commercial land use share	N/A	1%	41%	12%	13%	8%
Stops	7,427	2	1772	84	138	247
Vehicle-miles traveled	3,225,414	4,413	174,434	54,960	59,730	30,868
Jobs	286,143	2	61,550	2,936	5,299	9,647
Population per square mile	306,104	87	14,717	5,385	5,669	3,570
Under 18 population share	N/A	4%	27%	18%	18%	5%
Over 65 population share	N/A	2%	26%	10%	10%	4%
Combined walking and transit commute mode share	N/A	0%	26%	6%	7%	5%
Bicycle commute mode share	N/A	0%	3%	0%	1%	1%
Black/African-American population share	N/A	0%	31%	3%	4%	5%
Hispanic/Latino population share	N/A	9%	93%	48%	51%	23%
Poverty rate	N/A	2%	44%	16%	17%	10%

Metrolink Station Crash Predictors

Table 11 displays the results of the linear regression model for pedestrian collisions within ½ mile of Metrolink stations between 2013 and 2014. As the R^2 of 0.7342 indicates, approximately 73% of the variation in collision frequency across census tracts can be explained by the independent variables. However, similar to the Metro station model, only a few of the independent variables are statistically significant at the 0.1 level. The number of schools and people per square mile are significant at the 0.05 level, while the poverty rate is significant at the 0.1 level. Population density is the most influential predictor according to the standardized coefficients, while the poverty rate and the number of schools are also influential. As seen in the other models, the influence of the poverty rate suggests that high-poverty station areas have more people walking out of necessity, less safe infrastructure, and less safe behavior on the part of drivers. The influence of schools is more surprising, but schools may be one of the few pedestrian-generating land uses near Metrolink stations, which tend to be in relatively low-density suburban areas.

Table 12: Regression results for pedestrian-involved crashes by Metrolink station

Independent Variable	Unstandardized (b) Coefficient	Standard Error	t	P > |t|	Beta
Number of intersections	0.0054	0.0033	1.66	0.104	0.2540
Number of schools (natural log)	0.5440	0.2295	2.37	0.023	0.2572
Commercial land use share	0.0123	0.0139	0.88	0.383	0.0919
Major transit stops (natural log)	0.1444	0.1275	1.13	0.264	0.1510
Vehicle-miles traveled (natural log)	-0.0252	0.1982	-0.13	0.900	-0.0148
Jobs (natural log)	-0.1537	0.1101	-1.40	0.170	-0.1912
Population per square mile (natural log)	0.3795	0.1667	2.28	0.028	0.3541
Under 18 population share	-0.0424	0.0333	-1.27	0.210	-0.1821
Over 65 population share	0.0041	0.0262	0.16	0.875	0.0154
Combined walking and transit commute mode share	0.0117	0.0248	0.47	0.640	0.0506
Black/African-American population share	-0.0154	0.0262	-0.59	0.560	-0.0733
Hispanic/Latino population share	-0.0081	0.0081	-0.99	0.328	-0.1714
Poverty rate	0.0301	0.0171	1.76	0.087	0.2918
Number of observations	54	**R-squared**	0.7342	**Adjusted R-squared**	0.6478

Bicycle-Involved Crash Predictors

Table 12 displays the results of the linear regression model for bicycle-involved crashes within ½ mile of Metrolink stations. As the R^2 of 0.6704 indicates, approximately 67% of the variation between stations can be explained by this model. Similar to the pedestrian-involved crash model for Metrolink stations, few independent variables are statistically significant, with population per square mile the only variable significant at the 0.01 and 0.05 levels. The under 18 population share, which is negatively associated with bicycle crashes, is the only remaining significant predictor at the 0.1 level, although the poverty rate is close to that threshold with a P-value of 0.103. The standardized coefficients imply that population density, under 18 population share, and poverty rate are the most influential predictors of bicycle-involved crashes near Metrolink stations. The influence of population density is not surprising, but the crash-reduction influence of the under 18 population share suggests that young people and their families are less likely to bike or less likely to be injured while biking in areas near Metrolink stations. Additionally, the influence of poverty rate demonstrates that even when controlling for bicycle activity predictors, high-poverty communities are more likely to experience disproportionate collision rates.

Table 13: Regression results for bicycle-involved crashes by Metrolink station

Independent Variable	**Unstandardized (b) Coefficient**	**Standard Error**	**t**	**P > |t|**	**Beta**
Number of intersections	0.0019	0.0031	0.63	0.532	0.1071
Number of schools (natural log)	-0.0585	0.2164	-0.27	0.788	-0.0325
Commercial land use share	0.0068	0.0130	0.53	0.602	0.0602
Major transit stops (natural log)	-0.0072	0.1214	-0.06	0.953	-0.0089
Vehicle-miles traveled (natural log)	0.2645	0.1887	1.40	0.169	0.1830
Jobs (natural log)	-0.0719	0.1048	-0.69	0.496	-0.1052
Population per square mile (natural log)	0.5278	0.1614	3.27	0.002	0.5788
Under 18 population share	-0.0541	0.0314	-1.72	0.093	-0.2729
Over 65 population share	-0.0166	0.0248	-0.67	0.508	-0.0728
Bicycle commute mode share	0.2208	0.1479	1.49	0.143	0.1630
Black/African-American population share	-0.0165	0.0252	-0.66	0.515	-0.0924
Hispanic/Latino population share	-0.0014	0.0079	-0.18	0.858	-0.0355
Poverty rate	0.0268	0.0161	1.67	0.103	0.3056
Number of observations	54	**R-squared**	0.6704	**Adjusted R-squared**	0.5633

Policy Recommendations

Although this study does not explicitly test a policy or planning intervention, it has several implications for more equitable bicycle and pedestrian planning. At the tract level, the analysis clearly demonstrates an injustice in the SCAG region's transportation system, as high-poverty census tracts and predominately Hispanic or Latino tracts have endured a disproportionate share of the region's pedestrian- and bicycle-involved crashes between 2005 and 2014. Additionally, I found the number of major transit stops to be the strongest predictor of pedestrian-involved crashes and the third strongest predictor of bicycle-involved crashes by census tract. Additionally, at the Metro station level, I found vehicle-miles traveled to be a strong predictor of both crash types. My recommendations therefore suggest prioritizing the most vulnerable communities in the funding process and incentivizing safety enhancements near bus stops and rail stations. I also propose ways to strengthen and expand this study in the future.

California's Active Transportation Program (ATP) provides the bulk of pedestrian and bicycle safety funding statewide. Created in 2013 through the passage of Senate Bill 99 and Assembly Bill 101, ATP consolidated the existing Transportation Alternatives Program, Bicycle Transportation Account, and statewide Safe Routes to School programs to fund infrastructure, education, enforcement, and planning to encourage active transportation (Caltrans, 2018). Figure 14 shows the purpose and goals of ATP as illustrated by Caltrans (2015).

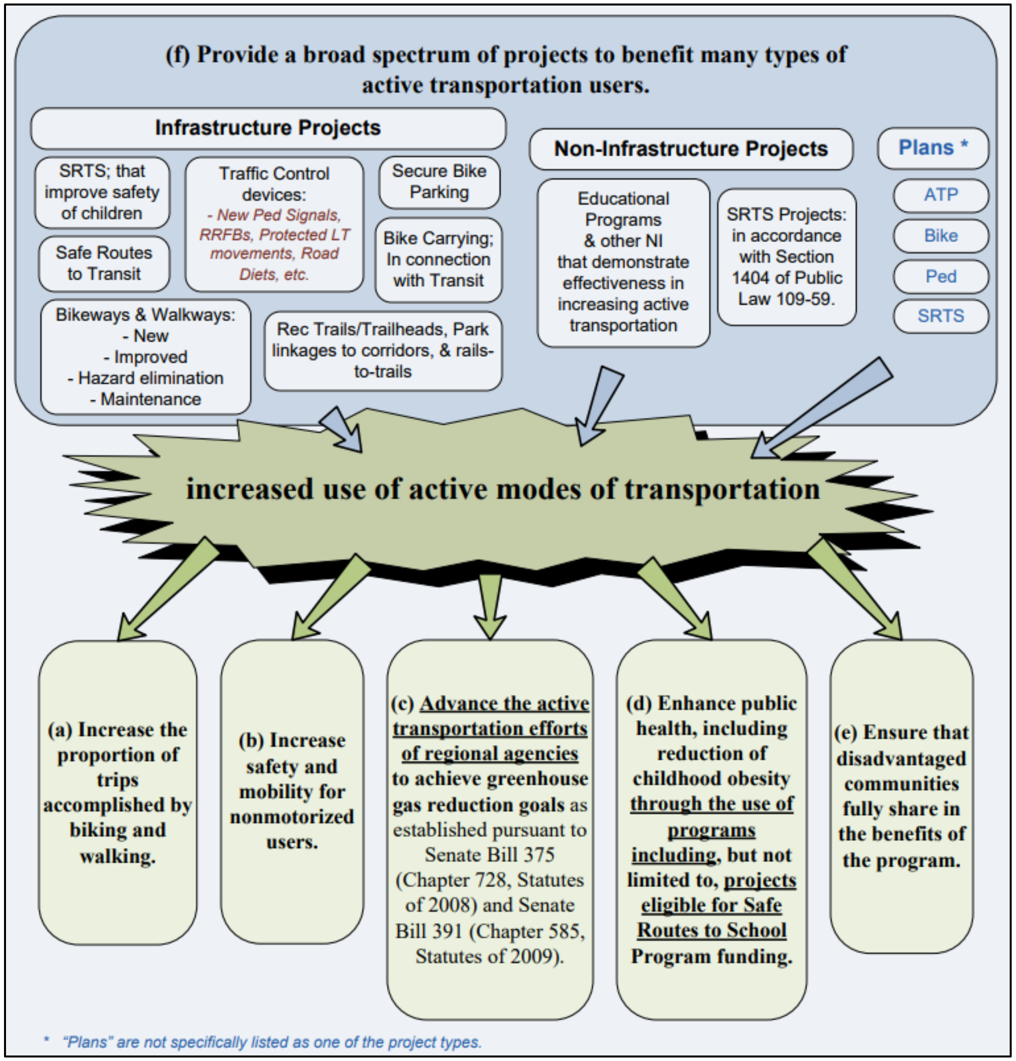

Figure 14: California Active Transportation Program purpose and goals. Source: Caltrans (2015)

Although the state government directs ATP fund allocation, metropolitan planning organizations (MPOs) including SCAG play an important role in the process. To obtain funds, public agencies must describe one or more projects in an application to the California Transportation Commission (CTC). The CTC scores each project and awards funds to those ranking highest until 50% of the ATP budget is used, with the caveat that 25% of program funds benefit disadvantaged communities (DACs) identified based on their median household income, exposure to pollution, and other factors (California Transportation Commission, 2018). Table 13 displays the CTC's scoring criteria.

Table 14: Statewide ATP scoring criteria. Source: California Transportation Commission (2018)

Scoring Topic	**Plan Application**	**Non-infrastructure only application**	**Infrastructure or Infrastructure/Non-infrastructure Applications**		
			Small	**Medium**	**Large**
Benefit to Disadvantaged Communities (DAC)	30	10	10	10	10
Need	20	40	53	43	38
Safety	N/A	10	25	25	20
Public Participation and Planning	25	15	10	10	10
Scope and Plan Consistency	N/A	10	2	2	2

Implementation and Plan Development	25	N/A	N/A	N/A	N/A
Context Sensitivity and Innovation	N/A	5	N/A	5	5
Transformative Projects	N/A		N/A	N/A	5
Evaluation and Sustainability	N/A	10	N/A	N/A	N/A
Cost Effective	N/A	N/A	N/A	N/A	5
Leveraging	N/A	N/A	N/A	5	5
Corps (0 or -5)	N/A	0	0	0	0
Past Performance (0 or -10)	0	0	0	0	0

The remaining projects are scored at the local or regional level, with 40% of ATP funds allocated to large MPOs including SCAG and 10% to jurisdictions outside of MPO territory. MPOs have the option to use the statewide evaluation criteria or develop their own application and scoring criteria, provided that 25% of the regional program's funds benefit DACs. Senate Bill 99, which created the ATP, also has unique requirements for SCAG, most notably that each county's transportation commissions must approve the final project list. For this reason, SCAG's regional guidelines for the 2017 ATP (the most recent year available) use the statewide scoring criteria, but with an additional

10 points available to each project, allocated at the discretion of the county transportation commissions. The resulting list is then modified, if necessary, to include the highest scoring projects benefitting DACs if the 25% threshold is not met in the original project list (Southern California Association of Governments, 2016b).

Prioritize Funding for Vulnerable Communities

This study reinforces the importance of the required 25% allocation to DACs and provides justification for expanding this requirement. As Lowe et al. (2015) demonstrate, a community's "equity advocacy capacity" often determines their likelihood of receiving public assistance for transportation projects. The 25% requirement acknowledges that DACs need additional support, particularly in the historically car-oriented SCAG region. Decades of overinvestment in high-speed automobile infrastructure are likely responsible for the unsafe walking and biking conditions in DACs, with the freeway-scarred Boyle Heights neighborhood as a particularly illustrative example (Breidenbach and Herrera, 2013). Environmental justice advocates emphasize the importance of rectifying these injustices by facilitating travel that is affordable and healthy, and this framework tends to favor active transportation investments (Creger et al., 2018). Unfortunately, according to local advocates, SCAG's largest county missed an opportunity to prioritize active transportation in the Measure M expenditure plan by dedicating less than 5% of funds to these improvements (IIP and LACBC, 2016).

Given the extent of active transportation needs in DACs, I recommend SCAG adopt stricter DAC funding requirements, ideally as part of the county-level supplemental project selection process. There are at least three ways SCAG could achieve this goal. One option would simply increase the required ratio of DAC to non-DAC projects from 25% to 35% or even 50%. Another option would maintain the existing 25% requirement but offer additional points to projects benefitting DACs. Finally, SCAG could maintain the 25% threshold while also requiring that 10% of the regional allocation go to the *most* disadvantaged communities as ranked by CalEnviroScreen. Finally, while it is important to make funds available to the lowest-income communities, it is equally important to ensure that the funding and implementation process is led by the community itself. Residents may be concerned about traffic impacts, higher property values leading to displacement, and facilities designed to serve higher-income newcomers in and around a neighborhood. Therefore, I recommend preserving or enhancing the points awarded for projects demonstrating thorough public participation and planning.

Enhance Safety near Major Transit Stops

As the number of major transit stops is a strong predictor of pedestrian- and bicycle-involved crashes at the census tract level, I recommend enhancing the ATP guidelines with language and scoring criteria pertaining to first-and-last mile connections to bus stops and rail stations. Transit riders are pedestrians for at least a small portion of their trip, and given the sprawling urban form of the SCAG region, most transit-dependent riders likely walk a substantial distance to the nearest stop. Transit stops therefore generate pedestrians and force them to walk on unsafe roads throughout the region. Furthermore, unreliable and infrequent service pressures riders to run to bus stops if they fear being late, encouraging unsafe behavior at intersections and driveways. SCAG should therefore award extra points to projects that facilitate safe access to transit stops, especially by enhancing visibility at crosswalks. As vehicle-miles traveled is a key indicator of crash risk near Metro rail stations, I also recommend awarding extra points for projects that would reduce average vehicle speeds near stations, which would mitigate risks posed by high VMT near major arterials.

Invest in Pedestrian and Bicycle Data Collection

A major limitation in this study is the lack of clear pedestrian and bicycle data for the SCAG region, and therefore I recommend that SCAG obtain existing data or invest in automatic counters. Although publicly available pedestrian and bicycle count data exists for some neighborhoods, it is sparse, covers very limited time periods, and difficult to compare to other count data. Automatic pedestrian and bicycle counters could provide a much more robust data source compared to what currently exists. Additionally, while the level of detail in the SWITRS data made this study possible, more detail on the survivors and victims of collisions, such as income, race, ethnicity, home location, and work location would allow researchers to study the impact of pedestrian and bicycle hazards on vulnerable communities more directly, rather than analyzing the demographics of the surrounding area. Therefore, I recommend advocating for the addition of standardized methods for collecting these data at crash locations.

Incorporate New Data in Future Research

In addition to bicycle and pedestrian activity data, other quantitative and qualitative data would enhance the impact of this project. Using the existing data, other statistical approaches such as negative-binomial regression could help verify the results of this study. Cluster analysis could also use the existing dataset to determine if certain combinations of predictor variables result in distinct crash rates. As my models already account for much of the spatial variation in crash frequency, incorporating data on sidewalk, bicycle facility, and pavement conditions could help determine the effectiveness of infrastructure improvements. Most importantly, researchers including Loukaitou-Sideris (2007) have emphasized the importance of "groundtruthing" quantitative studies with community-level input. Conducting field audits and participatory workshops with pedestrians and cyclists in the most vulnerable communities would provide a vital understanding of active transportation conditions in the region.

Acknowledgements

I am extremely thankful to the Research and Analysis Department at SCAG for hiring me as an intern in addition to serving as the client for this project. Jung Seo has been an incredibly accommodating supervisor, and Tom Vo generously offered to be my advisor on short notice, providing extensive feedback on the theoretical and technical aspects of the project. John Cho initially advised me as I designed an analysis of commercial gentrification, but was supportive when I changed my focus to traffic safety. Stephen Patchan from SCAG's Active Transportation team helped me understand the ATP funding process leading to my policy recommendations. Finally, I greatly benefitted from the camaraderie of my fellow SCAG interns, especially Sabrina Kim and Nicholas Maldarelli who joined me in facing the challenge of the UCLA MURP client project while working as R&A interns.

The Department of Urban Planning has also been incredibly supportive throughout this process. My faculty advisor, Chris Tilly, agreed to work with me on a study of commercial gentrification in relation to his economic development expertise, yet generously offered to remain my advisor after I changed my focus to traffic safety. Anastasia Loukaitou-Sideris, who partially inspired this study through her similar project in 2007, met with me multiple times and offered feedback on several drafts. Evelyn Blumenberg, who taught both quarters of our client project seminar, also met with me several times and helped me formulate cohesive versions of each section of this report. Professor Blumenberg's TA, Tam Guy, also offered insightful and detailed feedback on several sections, especially my literature review. Finally, I would not have had the GIS skills to even begin this project without the patient instruction of Leo Estrada, Norman Wong and Madeline Brozen in earlier GIS courses. I am also grateful for Madeline's feedback on my research design.

Last but not least, I am eternally grateful for the support of my closest friend and fellow MURP Roxana Aslan. Roxana supported me through many long nights and stressful days working on this project and helped me maintain a focused and positive attitude, all while completing her own coursework. I cannot imagine completing this without her help.

References

Blumenberg, E. A. (2017). Social Equity and Urban Transportation In G. Giuliano and S. Hanson (Ed.), *The Geography of Urban Transportation* (pp. 338-358). New York, New York: The Guilford Press.

Breidenbach, J., & Herrera, J. (2013). Learning in the Struggle: Reframing Economic & Environmental Justice. *Journal of Sustainability Education, 5,* 101-122

California Transportation Commission (2018). 2019 Active Transportation Program Guidelines. Retrieved from http://www.catc.ca.gov/programs/atp/2019/docs/051618_2019_ATP_Guidelines_Final_Adopted.pdf

Caltrans [California Department of Transportation]. (2018). Active Transportation Program. Retrieved from http://www.dot.ca.gov/hq/LocalPrograms/atp/

Caltrans [California Department of Transportation]. (2015). ATP Purpose and Goals Flowchart. Retrieved from http://www.dot.ca.gov/hq/LocalPrograms/atp/documents/2015/Visio-ATP-Purpose-and-Goals-flowchart-3-3-15.pdf

Chen, C., Lin, H., & Loo, B. P. Y. (2012). Exploring the Impacts of Safety Culture on Immigrants' Vulnerability in Non-motorized Crashes: A Cross-sectional Study. *Journal of Urban Health*, *89*(1), 138–152. https://doi.org/10.1007/s11524-011-9629-7

Cottrill, C. D., & Thakuriah, P. (Vonu). (2010). Evaluating pedestrian crashes in areas with high low-income or minority populations. *Accident Analysis & Prevention*, *42*(6), 1718–1728. https://doi.org/10.1016/J.AAP.2010.04.012

David, N. K.-B., & Rice, R. G. (1994). The role of the physical environment in child pedestrian accidents. *Journal of Advanced Transportation*, *28*(2), 171–187. https://doi.org/10.1002/atr.5670280206

Ewing, R., & Kreutzer, R. (2006). *Understanding the Relationship Between Public Health and the Built Environment*. Retrieved from http://www.activeliving.org/files/LEED_ND_report.pdf

Flanagan, E., Lachapelle, U., & El-Geneidy, A. (2015). Riding tandem: Does cycling infrastructure investment mirror gentrification and privilege in Portland, OR and Chicago, IL? *Research in Transportation Economics*, *60*, 14–24. https://doi.org/10.1016/j.retrec.2016.07.027

Graham, D. J., & Glaister, S. (2003). Spatial Variation in Road Pedestrian Casualties: The Role of Urban Scale, Density and Land-use Mix. *Urban Studies*, *40*(8), 1591–1607. https://doi.org/10.1080/0042098032000094441

IIP [Investing in Place] and LACBC [Los Angeles County Bicycle Coalition]. (2016). Analysis of Metro's Revised Expenditure Plan. Retrieved from https://investinginplace.org/wpcontent/uploads/2017/02/juneexpenditureplananalysis-lacbc-iip.pdf

Jacobsen, P. L. (2003). Safety in numbers: more walkers and bicyclists, safer walking and bicycling. *Injury Prevention*, *9*, 205–209. Retrieved from http://injuryprevention.bmj.com/content/injuryprev/9/3/205.full.pdf

Kim, K., Brunner, I., & Yamashita, E. (2006). Influence of Land Use, Population, Employment, and Economic Activity on Accidents. *Transportation Research Record: Journal of the Transportation Research Board*, *1953*, 56–64. https://doi.org/10.3141/1953-07

LaScala, E. A., Gerber, D., & Gruenewald, P. J. (2000). Demographic and environmental correlates of pedestrian injury collisions: a spatial analysis. *Accident Analysis & Prevention*, *32*(5), 651–658. https://doi.org/10.1016/S0001-4575(99)00100-1

Li, W., & Joh, K. (2016). Exploring the synergistic economic benefit of enhancing neighbourhood bikeability and public transit accessibility based on real estate sale transactions. *Urban Studies*, *54*(15), 42098016680147. https://doi.org/10.1177/0042098016680147

Litman, T. (2011). Distance-Based Vehicle Insurance Feasibility, Costs and Benefits. Retrieved from http://www.vtpi.org/dbvi_com.pdf

Loukaitou-Sideris, A., Liggett, R., & Sung, H.-G. (2007). Death on the Crosswalk. *Journal of Planning Education and Research*, *26*(3), 338–351. https://doi.org/10.1177/0739456X06297008

Lowe, K. (2016). Environmental Justice and Pedestrianism. *Transportation Research Record: Journal of the Transportation Research Board*, *2598*, 119–123. https://doi.org/10.3141/2598-14

Lubitow, A., & Miller, T. R. (2013). Contesting Sustainability: Bikes, Race, and Politics in Portlandia. *Environmental Justice*, *6*(4), 121–126. https://doi.org/10.1089/env.2013.0018

Lubitow, A., Zinschlag, B., & Rochester, N. (2016). Plans for pavement or for people? The politics of bike lanes on the “Paseo Boricua” in Chicago, Illinois. *Urban Studies*, *53*(12), 2637–2653. https://doi.org/10.1177/0042098015592823

Lugo, A. E. (2013). CicLAvia and human infrastructure in Los Angeles: ethnographic experiments in equitable bike planning. *Journal of Transport Geography*, *30*, 202–207. https://doi.org/10.1016/J.JTRANGEO.2013.04.010

McClintock, N. (2017). Cultivating (a) Sustainability Capital: Urban Agriculture, Ecogentrification, and the Uneven Valorization of Social Reproduction. *Annals of the American Association of Geographers*, 1–12. https://doi.org/10.1080/24694452.2017.1365582

National Research Council (2004). Understanding Racial and Ethnic Differences in Health in Late Life: A Research Agenda. *Washington (DC): National Academies Press (US).* 10, Health Care. Available from: https://www.ncbi.nlm.nih.gov/books/NBK24693/

Noland, R. B., Klein, N. J., & Tulach, N. K. (2013). Do lower income areas have more pedestrian casualties? *Accident Analysis & Prevention*, *59*, 337–345. https://doi.org/10.1016/J.AAP.2013.06.009

Organisation for Economic Co-operation and Development. (1998). Safety of Vulnerable Road Users. Retrieved from http://www.oecd.org/sti/transport/roadtransportresearch/2103492.pdf.

Pharr, J., Coughenhour, C., & Bungum, T. (2013, October 4). Environmental, Human and Socioeconomic Characteristics of Pedestrian Injury and Death in Las Vegas, NV. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2573539

Priyantha Wedagama, D. M., Bird, R. N., & Metcalfe, A. V. (2006). The influence of urban land-use on non-motorised transport casualties. *Accident Analysis & Prevention*, *38*(6), 1049–1057. https://doi.org/10.1016/J.AAP.2006.01.006

Retting, R. A., Ferguson, S. A., & McCartt, A. T. (2003). A review of evidence-based traffic engineering measures designed to reduce pedestrian-motor vehicle crashes. *American Journal of Public Health*, *93*(9), 1456–63. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12948963

Retting, R. A. (2018). Pedestrian Traffic Fatalities by State: 2017 Preliminary Data. Report prepared for the Governors Highway Safety Association, Washington, D.C. Retrieved from https://www.ghsa.org/sites/default/files/2018-03/pedestrians_18.pdf

Reynolds, C. C., Harris, M. A., Teschke, K., Cripton, P. A., & Winters, M. (2009). The impact of transportation infrastructure on bicycling injuries and crashes: a review of the literature. *Environmental Health*, *8*(1), 47. https://doi.org/10.1186/1476-069X-8-47

Rivara, F. P., & Barber, M. (1985). Demographic Analysis of Childhood Pedestrian Injuries. *Pediatrics*, *76*(3). Retrieved from http://pediatrics.aappublications.org/content/76/3/375

Sebert Kuhlmann, A. K., Brett, J., Thomas, D., & Sain, S. R. (2009). Environmental characteristics associated with pedestrian-motor vehicle collisions in Denver, Colorado. *American Journal of Public Health*, *99*(9), 1632–7. https://doi.org/10.2105/AJPH.2007.131961

Southern California Association of Governments. (2016a). The 2016-2040 Regional Transportation Plan / Sustainable Communities Strategy. Retrieved from http://scagrtpscs.net/Documents/2016/final/f2016RTPSCS.pdf.

Southern California Association of Governments. (2016b). 2017 Active Transportation Program Southern California Association of Government Regional Guidelines. Retrieved from http://www.catc.ca.gov/meetings/2016/2016-08/65_4.7.pdf

Stehlin, J. (2015). Cycles of Investment: Bicycle Infrastructure, Gentrification, and the Restructuring of the San Francisco Bay Area. *Environment and Planning A*, *47*(1), 121–137. https://doi.org/10.1068/a130098p

Tuckel, P., Milkarski, W., Maisel, R. (2014). Pedestrian injuries due to collisions with bicycles in New York and California. *Journal of Safety Research, 51,* 7-13. https://doi.org/10.1016/j.jsr.2014.07.003

Ukkusuri, S., Miranda-Moreno, L. F., Ramadurai, G., & Isa-Tavarez, J. (2012). The role of built environment on pedestrian crash frequency. *Safety Science*, *50*(4), 1141–1151. https://doi.org/10.1016/J.SSCI.2011.09.012

Wier, M., Weintraub, J., Humphreys, E. H., Seto, E., & Bhatia, R. (2009). An area-level model of vehicle-pedestrian injury collisions with implications for land use and transportation planning. *Accident Analysis & Prevention*, *41*(1), 137–145. https://doi.org/10.1016/J.AAP.2008.10.001

Yu, C.-Y. (2014). Environmental supports for walking/biking and traffic safety: Income and ethnicity disparities. *Preventive Medicine*, *67*, 12–16. https://doi.org/10.1016/J.YPMED.2014.06.028

Zhu, X., & Lee, C. (2008). Walkability and Safety Around Elementary Schools Economic and Ethnic Disparities. *Am J Prev Med*, *34*(4), 282–290. https://doi.org/10.1016/j.amepre.2008.01.024

Appendix

Appendix A: Tests of Multicollinearity

Multicollinearity by Census Tract

For my regression results to be valid, my independent variables must not exhibit a high amount of multicollinearity. As Table A1 shows, many of my independent variables are correlated with one another. The most correlated variable pairs at the census tract level are: population per square mile and combined walking and transit commute mode share (0.64) and poverty rate and combined walking and transit commute mode share (0.60). These correlations suggest some level of multicollinearity, yet a more systematic way to measure multicollinearity, known as the Variance Inflation Factor, reveals little cause for concern.

Table A1: Correlation matrix of independent variables at census tract level

	Inter - sections	Schools	Commer - cial land use share	Major transi t stops	Vehicle - miles travele d	Jobs	Popul - ation	Popul - ation per square mile	Under 18 popul - ation share	Over 65 popul - ation share	Black / African - America n popul - ation share	Hispanic / Latino popul - ation share	Pov - erty rate	Combined walking and transit commute mode share	Bicycle commute mode share
Inter - sections	1.00														
Schools	0.21	1.00													
Commer - cial land use share	-0.19	-0.09	1.00												

Transit stops	0.00	0.09	0.29	1.00											
Vehicle - miles traveled	0.57	0.16	-0.01	0.21	1.00										
Jobs	0.21	0.08	0.34	0.48	0.43	1.00									
Population	0.30	0.32	-0.09	0.02	0.21	0.10	1.00								
Population per square mile	-0.41	-0.16	0.36	0.03	-0.32	-0.14	-0.05	1.00							
Under 18 population share	-0.03	0.20	-0.18	-0.13	-0.04	-0.15	0.26	-0.01	1.00						
Over 65 population share	0.17	-0.09	-0.07	0.01	0.07	0.04	-0.21	-0.28	-0.49	1.00					
Black/ African - American population share	-0.08	0.02	0.01	0.11	-0.02	-0.04	0.00	0.08	0.09	-0.10	1.00				
Hispanic / Latino population share	-0.23	0.11	0.11	0.03	-0.11	-0.08	0.07	0.34	0.54	-0.48	-0.02	1.00			
Poverty rate	-0.16	0.03	0.27	0.14	-0.11	-0.02	-0.04	0.47	0.25	-0.32	0.21	0.56	1.00		
Combined walking and transit commute mode share	-0.26	-0.06	0.38	0.22	-0.15	0.03	-0.16	0.64	-0.08	-0.20	0.12	0.32	0.60	1.00	
Bicycle commute mode share	-0.11	-0.03	0.12	0.07	-0.07	0.03	-0.01	0.19	-0.11	-0.11	-0.04	0.08	0.20	0.26	1.00

The Variance Inflation Factor (VIF) measures the ratio of each variable’s variance alone to its variance in the regression model, and thus measures how multicollinearity impacts variance. Generally, a VIF greater than 5 suggests multicollinearity as a cause for concern. As Table A2 shows, VIF is less than 5 for all independent variables, with a mean VIF of 2.16. This suggests multicollinearity is not a problem in this analysis.

Table A2: Variance inflation factors of tract-level pedestrian crash regression model

Variable	VIF	1/VIF
lnint	4.34	0.23
lnpopdens	3.74	0.27
pctlatino	2.35	0.42
lnpctwalk	2.27	0.44
lntotalpop	2.25	0.44
lnjobs	2.22	0.45
pctbelowpov	2.06	0.49
pctunder18	2.05	0.49
lnpctcomm	1.89	0.53
pctover65	1.64	0.61
lnstops	1.63	0.61
lnvmt	1.42	0.70
lnschools	1.23	0.81
lnpctblack	1.18	0.84
Mean VIF	**2.16**	

Table A3: Variance inflation factors of tract-level bicycle crash regression model

Variable	**VIF**	**1/VIF**
Lnint	4.26	0.23
lnpopdens	3.81	0.26
pctlatino	2.37	0,42
lnpctbike	2.23	0.45
lntotalpop	2.29	0.44
lnjobs	2.03	0.49
pctbelowpov	2.17	0.46
pctunder18	2.09	0.48
lnpctcomm	1.74	0.57
pctover65	1.68	0.60
lnstops	1.85	0.54
lnvmt	1.38	0.72
lnschools	1.34	0.75
lnpctblack	1.29	0.78
Mean VIF	**2.18**	

Multicollinearity by Metro Station

Table A3 shows the correlation matrix for the variables included in the Metro station level analysis. It shows that pedestrian crash rates and bicycle crash rates are highly correlated (0.75). Relationships between independent variables are less strong, although the number of jobs is correlated closely with the number of transit stops (0.75) and commercial land use share (0.62). The poverty rate is also strongly correlated with the combined walking and transit commute mode share (0.77).

Table A4: Correlation matrix of independent variables at Metro station level

	Inter - sections	Schools	Commer - cial land use share	Major transit stops	Vehicle - miles traveled	Jobs	Popul - ation per square mile	Under 18 popul - ation share	Over 65 popul - ation share	Combined walking and transit commute mode share	Bicycle commut e mode share	Black / African - American popul - ation share	Hispa nic / Latino popul - ation share	Poverty rate
Inter - sections	1.00													
Schools	0.09	1.00												
Commer - cial land use share	0.15	-0.04	1.00											
Transit stops	0.28	0.18	0.40	1.00										
Vehicle - miles traveled	0.04	0.26	0.12	0.43	1.00									
Jobs	0.21	0.12	0.62	0.75	0.41	1.00								
Population per square mile	0.11	0.13	0.26	0.10	0.24	0.16	1.00							

Under 18 population share	0.10	0.11	-0.43	-0.36	-0.26	-0.49	0.11	1.00						
Over 65 population share	-0.19	-0.17	-0.00	0.21	0.17	0.23	-0.13	-0.18	1.00					
Combined walking and transit commute mode share	0.24	0.40	0.16	0.29	0.43	0.26	0.59	0.03	-0.22	1.00				
Bicycle commute mode share	0.01	-0.04	0.12	0.02	0.22	0.07	0.09	-0.14	-0.14	0.41	1.00			
Black / African – American population share	0.18	-0.05	-0.04	0.07	0.10	-0.07	-0.05	0.25	-0.01	-0.13	-0.18	1.00		
Hispanic / Latino population share	0.17	0.28	-0.33	-0.17	-0.07	-0.31	0.27	0.77	-0.34	0.39	0.06	-0.02	1.00	
Poverty rate	0.27	0.33	-0.05	0.14	0.24	0.01	0.36	0.35	-0.36	0.76	0.42	0.11	0.56	1.00

Despite these strong correlations, the VIF tests suggest that the Metro-level bicycle and pedestrian regression models are both valid. Figure A4 shows the VIF from the pedestrian crash regression model, revealing a maximum VIF of 5.23 and a mean VIF of 2.62. The VIFs for the bicycle regression model are generally lower, with a maximum of 4.99 and a mean of 2.26.

Table A5: Variance inflation factors of Metro station-level pedestrian crash regression model

Variable	**VIF**	**1/VIF**
intersections	1.31	0.76
lnpopdens	1.90	0.53
pctlatino	4.83	0.21
pctwalk	4.61	0.22
lnjobs	2.37	0.42
pctbelowpov	3.82	0.26
pctunder18	5.23	0.19
pctcomm	1.91	0.52
pctover65	1.69	0.59
lnstops	1.79	0.56
lnvmt	1.80	0.56
lnschools	1.28	0.78
pctblack	1.59	0.63
Mean VIF	**2.62**	

Table A6: Variance inflation factors of Metro station-level bicycle crash regression model

Variable	**VIF**	**1/VIF**
intersections	1.32	0.76
lnpopdens	1.83	0.55
pctlatino	4.52	0.22
Pctbike	1.65	0.61
lnjobs	2.28	0.44
pctbelowpov	2.77	0.36
pctunder18	4.99	0.20
pctcomm	1.89	0.53
pctover65	1.70	0.59
lnstops	1.88	0.53
lnvmt	1.72	0.58
lnschools	1.31	0.77
pctblack	1.57	0.64
Mean VIF	**2.26**	

Metrolink Station

Table A5 shows the correlation matrix for the variables included in the Metrolink station level analysis. Among independent variables, population density is strongly correlated with the number of intersections (0.71), while the Hispanic/Latino population share is strongly correlated with the under 18 population share (0.72) and the poverty rate (0.66).

Table A7: Correlation matrix of independent variables at Metrolink station level

	Inter - sections	Schools	Commer - cial land use share	Major transi t stops	Vehicle - miles travele d	Jobs	Popul - ation per square mile	Under 18 popul - ation share	Over 65 popul - ation share	Combin ed walking and transit commu te mode share	Bicycle commu te mode share	Black / African - Americ an popul - ation share	Hispanic / Latino popul - ation share	Pov - erty rate
Inter - sections	1.00													
Schools	0.45	1.00												
Commer - cial land use share	0.05	-0.03	1.00											
Transit stops	0.39	0.20	0.05	1.00										
Vehicle - miles traveled	0.36	0.04	0.30	0.60	1.00									
Jobs	0.48	0.20	0.43	0.47	0.47	1.00								

Population per square mile	0.71	0.53	0.06	0.44	0.29	0.36	1.00							
Under 18 population share	0.13	0.22	-0.09	-0.39	-0.20	-0.20	0.11	1.00						
Over 65 population share	0.16	-0.02	-0.10	0.13	0.01	-0.05	0.12	-0.31	1.00					
Combined walking and transit commute mode share	0.48	0.38	-0.03	0.47	0.38	0.42	0.46	-0.11	0.05	1.00				
Bicycle commute mode share	0.34	0.39	0.40	0.27	-0.02	0.09	0.10	0.28	0.46	0.02	1.00			
Black/ African - American population share	0.12	-0.18	-0.00	0.14	0.11	0.17	0.01	-0.07	0.17	0.10	-0.07	1.00		
Hispanic / Latino population share	0.24	0.27	-0.01	0.09	0.13	0.11	0.26	-0.07	0.72	-0.33	0.13	-0.10	1.00	
Poverty rate	0.40	0.35	-0.05	0.13	0.12	0.12	0.34	0.53	-0.18	0.28	0.04	0.37	0.66	1.00

These correlations suggest that the Metrolink-level variables exhibit more multicollinearity than the tract- and Metro-level variables, and the VIF tests confirm that. Nonetheless, the maximum VIFs of 4.52 for the pedestrian crash regression and 4.72 for the bicycle crash regression shown in Table A6 do not exceed the threshold of 10, implying that both models are valid.

Table A8: Variance inflation factors of Metrolink station-level pedestrian crash regression model

Variable	**VIF**	**1/VIF**
intersections	3.51	0.28
lnpopdens	3.64	0.27
pctlatino	4.52	0.22
pctwalktrans	1.74	0.58
lnjobs	2.82	0.35
pctbelowpov	4.16	0.24
pctunder18	3.08	0.32
pctcomm	1.63	0.61
pctover65	1.44	0.70
lnstops	2.68	0.37
lnvmt	2.05	0.49
lnschools	1.77	0.56
pctblack	2.34	0.43
Mean VIF	**2.72**	

Table A9: Variance inflation factors of Metrolink station-level bicycle crash regression model

Variable	**VIF**	**1/VIF**
Intersections	3.50	0.29
lnpopdens	3.80	0.26
pctlatino	4.72	0.21
pctbike	1.45	0.69
lnjobs	2.85	0.35
pctbelowpov	4.06	0.25
pctunder18	3.04	0.33
pctcomm	1.59	0.63
pctover65	1.44	0.70
lnstops	2.70	0.37
lnvmt	2.07	0.48
lnschools	1.75	0.57
lnpctblack	2.41	0.42
Mean VIF	**2.72**	

Appendix B: Crashes by Station

Pedestrian Crashes by Metro Station

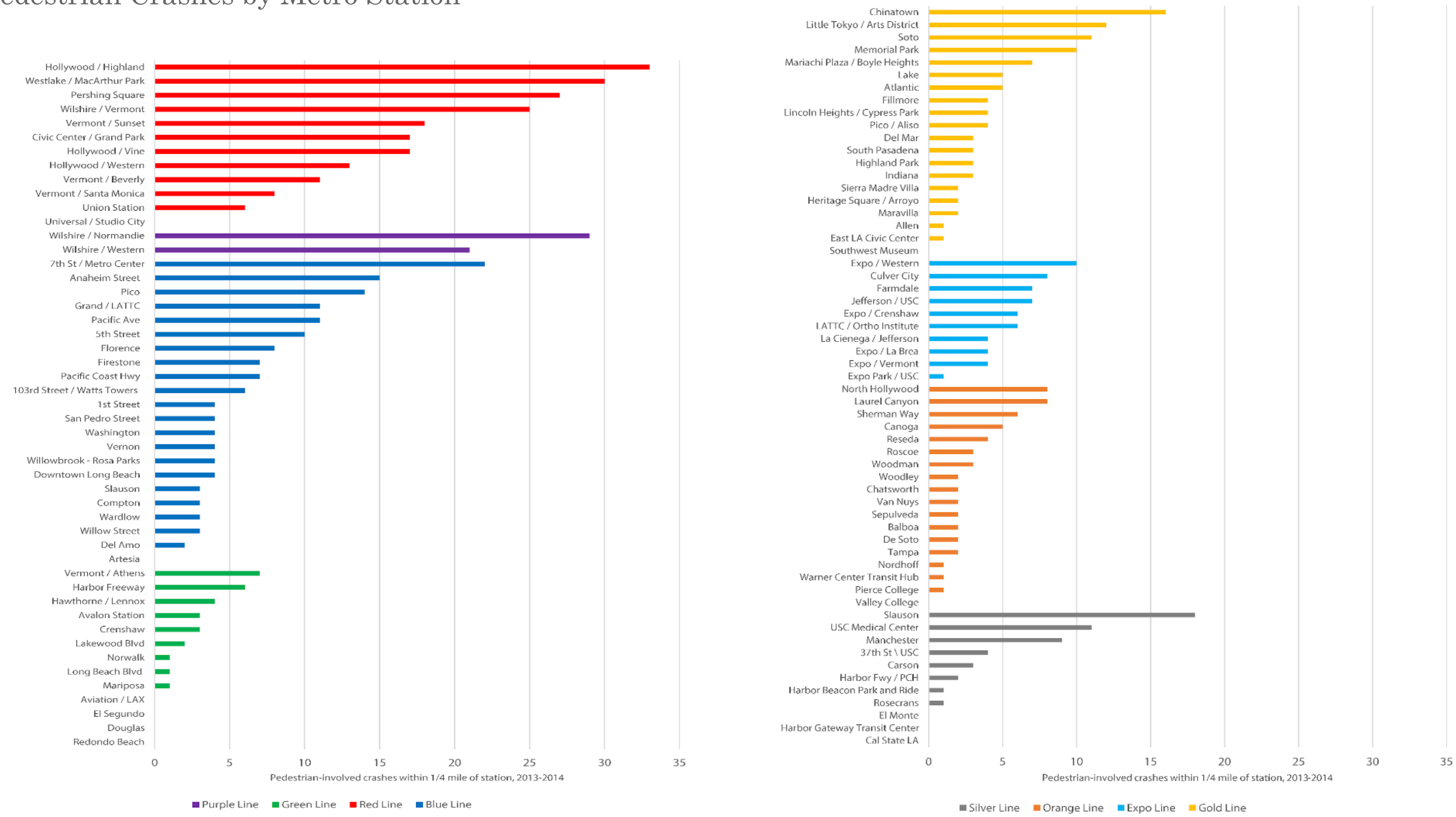

Figure B1: Pedestrian-involved crashes within ¼ mile of Metro station, 2013-2014

Bicycle-Involved Crashes by Metro Station

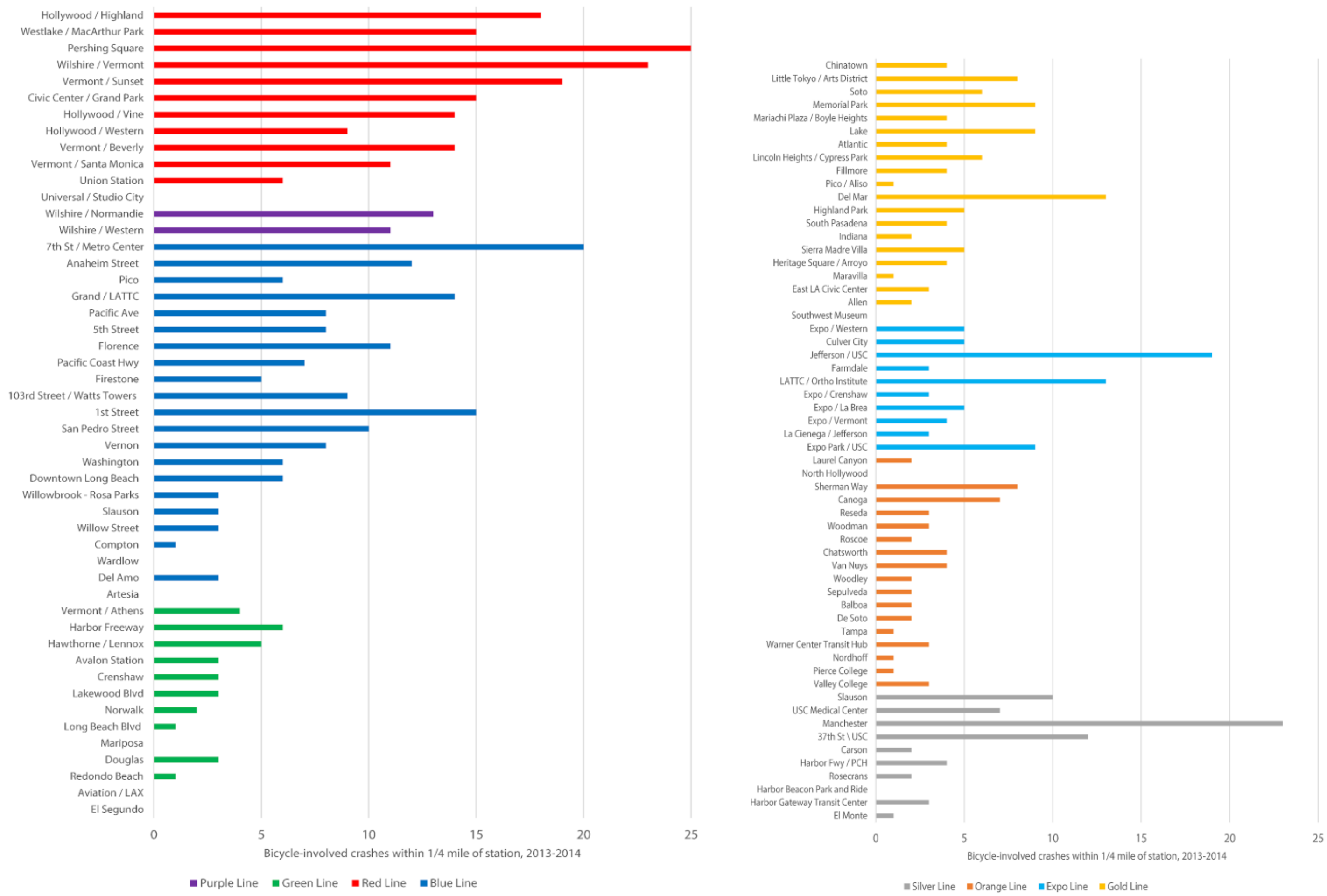

Figure B2: Bicycle-involved crashes within ¼ mile of Metro station, 2013-2014

Pedestrian-Involved Crashes by Metrolink Station

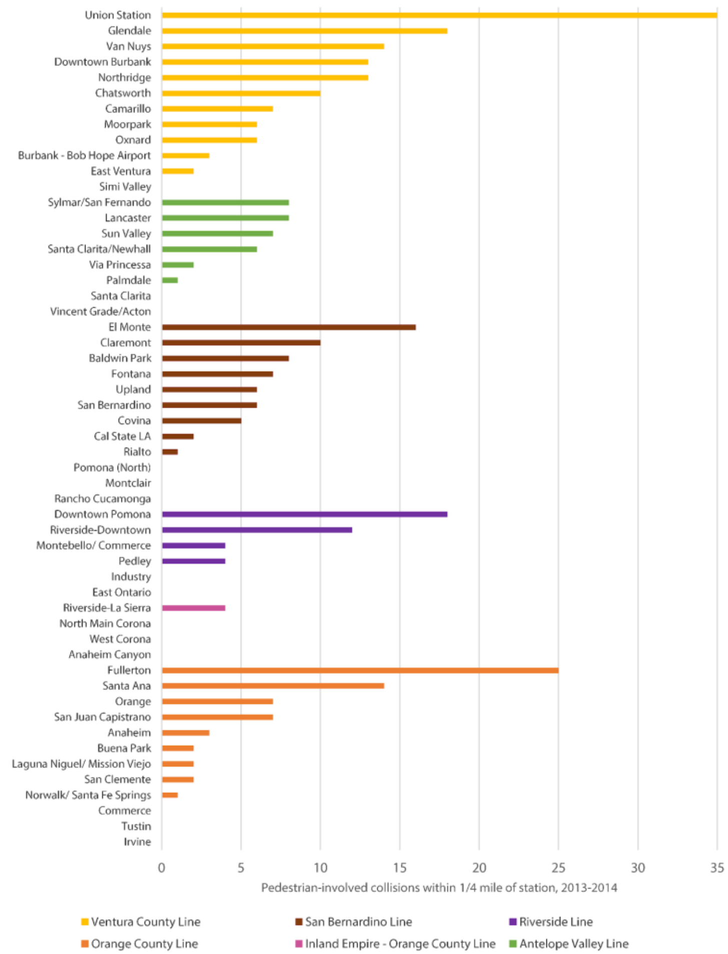

Figure B3: Pedestrian-involved crashes within ½ mile of Metrolink station, 2013-2014

Bicycle-Involved Crashes by Metrolink Station

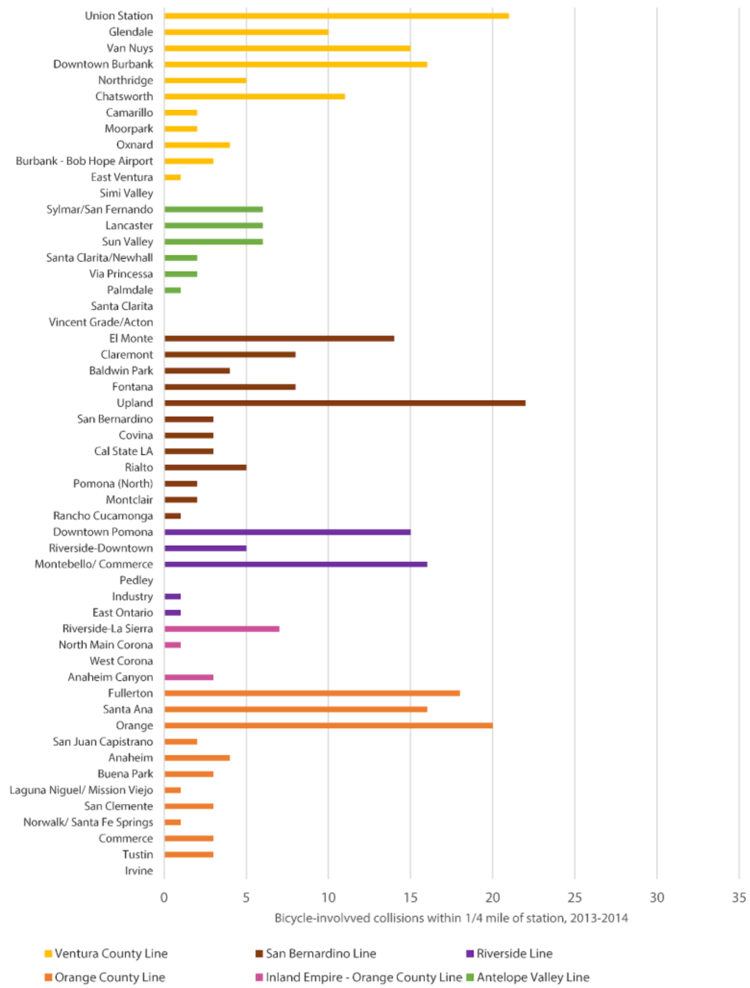

Figure B4: Bicycle-involved crashes within ½ mile of Metrolink station, 2013-2014