

Race, Class, and the Production of and Exposure to Vehicular Pollution in Los Angeles

July
2021

A Research Report from the Pacific Southwest
Region University Transportation Center

Geoff Boeing, University of Southern California
Yougeng Lu, University of Southern California
Clemens Pilgram, University of Southern California
Peter Mannino, University of Southern California



TECHNICAL REPORT DOCUMENTATION PAGE

1. Report No. PSR-20-19	2. Government Accession No. N/A	3. Recipient's Catalog No. N/A	
4. Title and Subtitle Race, Class, and the Production of and Exposure to Vehicular Pollution in Los Angeles		5. Report Date July 2021	
		6. Performing Organization Code N/A	
7. Author(s) Geoff Boeing, 0000-0003-1851-6411 Yougeng Lu Clemens Pilgram Peter Mannino		8. Performing Organization Report No. PSR-20-19	
9. Performing Organization Name and Address METRANS Transportation Consortium University of Southern California 650 Childs Way, RGL 216 Los Angeles, CA 90089		10. Work Unit No. N/A	
		11. Contract or Grant No. USDOT Grant 69A3551747109	
12. Sponsoring Agency Name and Address U.S. Department of Transportation Office of the Assistant Secretary for Research and Technology 1200 New Jersey Avenue, SE, Washington, DC 20590		13. Type of Report and Period Covered Final report (Aug. 2020 – Aug 2021)	
		14. Sponsoring Agency Code USDOT OST-R	
15. Supplementary Notes Project page URL: https://metrans.org/research/race-class-and-the-production-of-and-exposure-to-vehicular-pollution-in-los-angeles-			
16. Abstract Vehicular air pollution has created an ongoing public health crisis. Despite growing knowledge of racial injustice in exposure to vehicular particulate matter (PM2.5), less is known about the sociodemographics-mediated relationship between the production of and exposure to such pollution. This study assesses pollution burden with a unified indicator measuring local populations' exposure to PM2.5 adjusted by their own vehicle kilometers traveled. Through a Los Angeles case study we examine how production-adjusted exposure to vehicular PM2.5 relates to race/ethnicity and socioeconomic status, and how this relationship varies across the region. We find that, all else equal, tracts whose residents drive less experience more air pollution, as do tracts with a more non-White population. Our commute simulation demonstrates how commuters from majority- White tracts disproportionately travel through majority non-White areas to drive to work. Decades of racially-motivated freeway infrastructure planning and residential segregation shape today's disparities between who produces vehicular pollution and who is exposed to it. We conclude by suggesting paths toward racial justice at the nexus of urban transport and environmental planning, and discussing various policy interventions.			
17. Key Words Air pollution, vehicular pollution, social justice, inequality, sociodemographic		18. Distribution Statement No restrictions.	
19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 34	22. Price N/A

Contents

About the Pacific Southwest Region University Transportation Center	3
U.S. Department of Transportation (USDOT) Disclaimer	3
Disclosure	3
Acknowledgements	4
Abstract	5
1 Introduction	8
2 Background	9
2.1 Transport and Emissions	9
2.2 Emissions and Health	9
2.3 Race and Class Disparities	10
2.4 Open Problem	11
3 Methods	11
3.1 Input Data	11
3.1.1 Air Pollution	13
3.1.2 Passenger Travel	13
3.1.3 Freight Travel	15
3.1.4 Street Network	15
3.1.5 Demographics	15
3.2 Modeling and Analysis	15
3.2.1 OLS Models	15
3.2.2 GWR Models	16
3.2.3 Commute Simulation	17
4 Results	18
4.1 Descriptive Statistics	18
4.2 OLS Results	18
4.3 GWR Results	22
4.4 Commute Simulation Results	22
5 Discussion	24
6 Conclusion	29

About the Pacific Southwest Region University Transportation Center

The Pacific Southwest Region University Transportation Center (UTC) is the Region 9 University Transportation Center funded under the US Department of Transportation's University Transportation Centers Program. Established in 2016, the Pacific Southwest Region UTC (PSR) is led by the University of Southern California and includes seven partners: Long Beach State University; University of California, Davis; University of California, Irvine; University of California, Los Angeles; University of Hawaii; Northern Arizona University; Pima Community College. The Pacific Southwest Region UTC conducts an integrated, multidisciplinary program of research, education and technology transfer aimed at improving the mobility of people and goods throughout the region. Our program is organized around four themes: 1) technology to address transportation problems and improve mobility; 2) improving mobility for vulnerable populations; 3) Improving resilience and protecting the environment; and 4) managing mobility in high growth areas.

U.S. Department of Transportation (USDOT) Disclaimer

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated in the interest of information exchange. The report is funded, partially or entirely, by a grant from the U.S. Department of Transportation's University Transportation Centers Program. However, the U.S. Government assumes no liability for the contents or use thereof.

Disclosure

Principal Investigator, Co-Principal Investigators, others, conducted this research titled, "Title of Project" at [Department, School, University]. The research took place from [start date] to [end date] and was funded by a grant from the [funding source] in the amount of [\$amount]. The research was conducted as part of the Pacific Southwest Region University Transportation Center research program.

Acknowledgements

This research was supported by a grant from the Pacific Southwest Region University Transportation Center. The authors wish to thank Nicholas Cerdera and David Rosas Flores for research assistance.

Abstract

Vehicular air pollution has created an ongoing public health crisis. Despite growing knowledge of racial injustice in exposure to vehicular particulate matter (PM_{2.5}), less is known about the sociodemographics-mediated relationship between the production of and exposure to such pollution. This study assesses pollution burden with a unified indicator measuring local populations' exposure to PM_{2.5} adjusted by their own vehicle kilometers traveled. Through a Los Angeles case study we examine how production-adjusted exposure to vehicular PM_{2.5} relates to race/ethnicity and socioeconomic status, and how this relationship varies across the region. We find that, all else equal, tracts whose residents drive less experience more air pollution, as do tracts with a more non-White population. Our commute simulation demonstrates how commuters from majority-White tracts disproportionately travel through majority non-White areas to drive to work. Decades of racially-motivated freeway infrastructure planning and residential segregation shape today's disparities between who produces vehicular pollution and who is exposed to it. We conclude by suggesting paths toward racial justice at the nexus of urban transport and environmental planning, and discussing various policy interventions.

Race, Class, and the Production of and Exposure to Vehicular Pollution in Los Angeles

Executive Summary

Exposure to air pollution from cars, including particulate matter of 2.5 microns or smaller ($PM_{2.5}$), poses a significant risk to human health. In US cities, individuals' exposure is largely determined by residential locations, climate and terrain conditions, travel behavior, and racially-motivated planning decisions. For over a century, planners and politicians have both intentionally and unintentionally segregated housing markets, incentivized white flight, and bulldozed predominantly minority and lower-income neighborhoods to build freeways linking higher socioeconomic-status suburbs to job centers. These planning decisions manifest themselves today in disparate transportation infrastructure usage and driving behavior that concentrate hazardous vehicular pollution in minority communities. But how much vehicular pollution are different communities exposed to in relation to how much—and where—their residents drive?

This study examines this question in the context of race, class, and driving in Los Angeles County. Responding to growing knowledge of racial and socioeconomic injustice in exposure to vehicular $PM_{2.5}$, we measure the association between local populations' exposure to $PM_{2.5}$ and their own vehicle kilometers traveled (VKT). Prior research has exposed racial and socioeconomic inequities in exposure to pollution, or documented racially-motivated freeway infrastructure planning—but does not explicitly evaluate the spatial mismatch between the home locations of freeway users and where the emissions resulting from their driving end up. We do so with a Los Angeles case study that takes a twofold analytical approach, using ordinary least squares (OLS) and geographically weighted regression (GWR) regression analysis to measure the relationship between driving and emissions exposure, and using a commute simulation to identify who generates the excess driving through areas that rank lower in terms of VKT production.

In our regression analyses, we model census tract-level $PM_{2.5}$ concentrations attributable to on-road sources as a function of VKT by tract residents, controlling for local demographics, socioeconomic status, street network characteristics. Relying on tract-level VKT production data from Local Area Transportation Characteristics for Households (LATCH) and vehicular $PM_{2.5}$ emissions data from the Union of Concerned Scientists, we find a negative association between VKT and exposure to vehicular $PM_{2.5}$: all else equal, residents of communities that drive more tend to be less exposed to vehicular pollution themselves. A 1% increase in VKT production is associated with a 0.62% decrease in local $PM_{2.5}$ exposure. Allowing for heterogeneity in this statistical relationship by using GWR instead of OLS, we find substantial variation across the county, but the association between VKT and $PM_{2.5}$ exposure is negative in far more tracts than it is positive.

A second analysis evaluates the routes chosen respectively by white and by non-white commuters within Los Angeles County to understand how the demographics of commuters traversing a tract resemble the tract's residents and to identify where they most differ. Using census block-level home and work locations, we simulate commute routes and compare the share of all commuting meters driven through a tract by non-Hispanic white commuters to the share of the tract's population

that is non-Hispanic white. While commutes originating in majority-white tracts often traverse great distances through majority non-white parts of the county, the reverse is far less common: Los Angeles County is only 26% non-Hispanic white but, along major regional freeways, we observe discrepancies of up to 13 percentage points between the shares of traversing commuters and of the local population that are white. By utilizing freeways, White commuters traverse tracts that are far more non-White than the tracts where most of them live, on average. They disproportionately receive the benefits of driving on a highway, but because those highways are predominantly in non-White neighborhoods, other racial groups bear many of the external costs of that driving.

To return to our original question, we find overall that different communities are not exposed to vehicular pollution at a level proportional to how much they drive, and past infrastructure planning and housing policy play a central role in explaining these disparities. All else equal, tracts whose residents drive less experience more air pollution. Furthermore, tracts with a larger non-White population share—whether high income or low income—experience more air pollution than do Whiter but otherwise similar tracts. On average, White commuters traverse tracts that are far more non-White than the tracts where they live, but non-White commuters do not travel through tracts that are substantially Whiter than their own. This reveals an injustice in pollution burden with a distinct racial dimension, as the burden of pollution exposure falls disproportionately on communities that drive less and that are on average less white. This is largely driven by past planning and policy, and it is up to current planners and policymakers to attenuate this injustice—but there is no silver bullet. Nevertheless, vehicle electrification, congestion tolls, improved public transit options, and better housing policy may help mitigate these harms.

1 Introduction

Throughout the mid-twentieth century, planners routed and built an enormous network of expressways to open up the growing Los Angeles metropolis to motorists. Vast swaths of established urban neighborhoods were bulldozed to clear new channels for suburban residents to drive to job centers. Yet some older neighborhoods survived relatively unscathed. For example, local residents organized to protest, and eventually successfully cancel, plans to extend State Route 2 through the affluent communities of Beverly Hills and Los Angeles's westside. In contrast, similar grassroots efforts failed in Los Angeles's eastside, where several major freeways carved up its less-affluent and less-white neighborhoods. This combination of race, affluence, and political power shaped infrastructure planning that today determines regional accessibility, travel behavior, and pollution exposure.

Exposure to air pollution from cars, particularly particulate matter of 2.5 microns or smaller (PM_{2.5}), poses a significant risk to human health (Alman et al., 2019; Anderson et al., 2012; Bell and Ebisu, 2012; Ljubimova et al., 2018; Loftus et al., 2020). Much of the transportation-environmental justice literature around air pollution emphasizes community disparities in exposure, focusing on residential proximity to toxic release sites, refineries, road infrastructure, etc (Bravo et al., 2016; Mikati et al., 2018; Schweitzer and Valenzuela, 2004; Reichmuth, 2019a; Jorgenson et al., 2020; Liévanos, 2019). These justice claims are straightforward when identifying a non-mobile emissions source, like a refinery: the refinery exposes nearby residents to harmful pollutants through a spatial diffusion process.

However, mobile emissions sources, such as cars, are less straightforward. When justice arguments rely on residential proximity to highway infrastructure, they identify the highway as the emissions source. In reality, of course, it is vehicles utilizing the highway that emit pollution. The Racially-motivated political/planning decisions behind transportation infrastructure placement impact environmental justice today through vehicular pollution. Importantly, the polluting vehicles may be driven by people with hypothetically different sociodemographic profiles than the communities they traverse and pollute. Despite our growing knowledge of pollution exposure injustice, we know less about the sociodemographics-mediated relationship between who generates pollution and who is exposed to that pollution. The individuals exposed to the most vehicular pollution may simply be producing the most themselves. A better measure of environmental justice would ask: how much vehicular pollution are you exposed to in relation to how much you drive?

This study examines this question in the context of race, class, and driving in Los Angeles. We develop and analyze a new indicator, the Burden Ratio (BR), of local exposure to vehicular pollution per vehicle kilometer traveled (VKT). We then model BR as a function of race, income, and other covariates to better understand disparate infrastructure use and impacts in Los Angeles. Using ordinary least squares (OLS) and geographically-weighted regression (GWR) we find that, all else equal, tracts that generate more vehicular travel tend to be exposed to less vehicular air pollution. Minority and poorer communities are disproportionately burdened with excess vehicular pollution. Our GWR models allow us to explain substantially more of the variation than the OLS models that are standard in the environmental justice literature. Moving beyond the original placement of infrastructure, we probe racial justice and equity dimensions to suggest paths forward, including

identifying specific communities and possible interventions such as VKT taxes, tolls, emissions standards, and electrification.

This paper is organized as follows. It begins with a literature review to situate this study. Then it introduces our data sources and analytical methods. Next it presents our findings. Finally it concludes with a discussion of transportation, environmental justice, and interventions for planners and policymakers.

2 Background

Vehicular transportation generates urban air pollution that can harm human health. This relationship is mediated by the historical placement and current utilization of transportation infrastructure by motorists. This section reviews the literature on transport, emissions, health, and race/class disparities to situate the present study.

2.1 Transport and Emissions

Many studies have identified environmental externalities stemming from urban transportation including greenhouse gas emissions and air, water, and noise pollution. Among these, air pollution has received the most attention (Button, 1990). Delucchi (2000) estimated motor vehicles' environmental externalities and found that, compared to other environmental hazards, air pollution's costs are the greatest—particularly for human health. Researchers have identified cars and trucks as the most significant source of urban air pollution (Karner et al., 2010; Rowangould, 2015; Pan et al., 2013) and several recent studies have quantified the air pollution gradient near highways (Brugge et al., 2007; Seagram et al., 2019; Zhu et al., 2002).

The US Environmental Protection Agency (EPA) reports that transportation accounts for significant carbon monoxide, nitrogen oxides, PM_{2.5}; sulfur oxide, and volatile organic compounds emissions (US EPA, 2018). In California, transportation generates 37% of the state's GHG emissions—more than any other sector (Board, 2020). In Southern California, vehicular emissions generate roughly a third of the region's PM_{2.5} (Habre et al., 2021; Hasheminassab et al., 2014). Truck freight is an important source. Brunekreef et al. (1997) found that truck traffic can lead to more pollution-related health impacts than personal automobile traffic. These externalities concentrate in freight hubs and transportation corridors (Wu et al., 2009).

2.2 Emissions and Health

Air pollution is associated with an increased risk of lung cancer, asthma, bronchitis, and impaired cognitive development. The research literature documents numerous adverse health impacts from vehicular air pollution such as particulate matter, carbon monoxide, and ozone. This body of evidence arises from studies assessing proximity to highways, direct exposure to air pollutants, or both.

Air pollution can cause acute respiratory symptoms, infant bronchitis, and decreased lung growth in children. Ambient particulate matter and gaseous pollutants are associated with increased

childhood asthma hospitalization (Lee et al., 2006). Yin et al. (2011) found that higher daily PM_{2.5} concentrations correspond to higher pediatric respiratory rates. Several studies have demonstrated that children exposed to high-traffic roadways suffer from more respiratory ailments like asthma (Gauderman et al., 2005; Janssen et al., 2001; Oosterlee et al., 1996). Exposure to vehicular emissions can increase school absence, cause asthma, and exacerbate existing asthma (Künzli et al., 2003). A UK hospital admission survey revealed that asthma patients were more likely to live near high-traffic roadways (Beris and Edwards, 1994), and a Swedish study found that exposure to NO₂ can cause bronchitis (Pershagen et al., 1995).

Beyond these childhood respiratory ailments, epidemiologists have implicated airborne particulate matter in several other health impacts. Air pollution exposure is associated with lung cancer incidence, especially among people who have never smoked (Beelen et al., 2008; Jagai et al., 2017). WHO guidelines indicate that when PM_{2.5} concentrations reach 50 µg/m³, short-term mortality increases 2.5%, mainly attributed to lung cancer (World Health Organization, 2006). Researchers have also established a link between vehicular air pollution and cognitive development: children exposed to more air pollution demonstrate slower cognitive development (Sunyer et al., 2015).

2.3 Race and Class Disparities

These health impacts pose an urgent problem, but one that is unevenly distributed across populations. Environmental hazards disproportionately burden minorities and low-income individuals relative to whiter and more affluent ones (Bae et al., 2007; Liévanos, 2019; Houston et al., 2004, 2008; Tessum et al., 2021). These vulnerable populations often have limited residential mobility and reside in undesirable locations near freeways (Bae et al., 2007). Rowangould (2013) finds that in many US cities, neighborhoods within 500 meters of roads with at least 25,000 average daily vehicle trips are majority Latino or Black. Furthermore, freight traffic concentrates in minority and low-income neighborhoods adjacent to ports (Houston et al., 2008). Many studies have documented that minority and low-income communities are disproportionately burdened by transportation externalities while simultaneously suffering from worse accessibility than whiter, more affluent communities (Rowangould, 2015; Schweitzer and Valenzuela, 2004; Rowangould et al., 2016; Yuan, 2018; Houston et al., 2014; Poorfakhræi et al., 2017; Pratt et al., 2015). In tandem, this suggests that poor and minority communities do not benefit from transportation infrastructure in proportion to the costs they bear.

The confluence of transportation, health, and racial injustice is particularly extreme in Los Angeles, a poster child for automobile dependence, air pollution, racial diversity, and wealth inequality. In 2019, 74% of Los Angeles County residents were nonwhite. Latinos¹ (49% of total population) are the single largest group followed by Whites² (26%), Asians (15%) and Blacks³ (9%)⁴. The region's poverty rate is consistently higher than the national average. In Los Angeles,

¹“Latino” is used throughout this paper as shorthand for Hispanic or Latino of any race

²“White” is used throughout this paper as shorthand for non-Hispanic or Latino White alone

³“Black” is used throughout this paper as shorthand for non-Hispanic or Latino Black alone

⁴U.S. Census Bureau QuickFact: Los Angeles County, California (see <https://www.census.gov/quickfacts/fact/table/losangelescountycalifornia,CA/PST045219>)

minority and low-income populations are more likely to live in central neighborhoods for better accessibility and are exposed to twice the local traffic density as the rest of the region (Giuliano, 2003; Houston et al., 2004). Despite this exposure, they are less likely to own an automobile and more likely to commute by public transit (PolicyLink and USC Program for Environmental and Regional Equity, 2017).

Decades of racist planning decisions in Los Angeles have caused today's injustices in transportation, health, and environmental quality. In the 20th century, transportation, housing, and land use planning accommodated white flight and industry at the expense of minority communities (Estrada, 2005; Baum-Snow, 2007). Planners bulldozed central neighborhoods to construct an expansive freeway system and create regional access for new peripheral suburbs and their predominately White residents (Pulido, 2000), usually selecting freeway alignments through minority neighborhoods for the sake of "slum clearance" (Mohl, 2004; DiMento, 2009; White House, 2021) or to protect business interests (Perez, 2017). Freeway construction fragmented and displaced these minority communities and subjected nearby residents to decades of resulting air pollution (Mohl, 2004). While some freeways were slated to be built through whiter and more affluent neighborhoods, they were often cancelled or rerouted due to the neighborhoods' political clout, as was the case with the California State Route 2 through Beverly Hills (Perez, 2017; Masters, 2014).

2.4 Open Problem

Schweitzer and Valenzuela (2004) argue that transportation justice fundamentally concerns the distribution of transportation costs and benefits. The literature suggests that some groups disproportionately benefit from automobility while other groups disproportionately bear its costs. The present study advances this literature through a Los Angeles case study investigating the relative production of and exposure to vehicular pollution. Are different communities exposed to vehicular pollution at a level proportional to how much they drive? If not, what is the ceteris paribus relationship between race and this disparity? This knowledge is essential for guiding planning efforts, targeting policy interventions, and setting specific environmental justice goals.

3 Methods

3.1 Input Data

This study focuses on Los Angeles County (Figure 1). Urban scholars often identify Los Angeles as the poster child of turn-of-the-millennium urbanization (e.g. Dear, 2001) and its sprawling structure, transportation infrastructure planning, and residential segregation allow for insights generalizable to other American cities. We collect data on air pollution, passenger vehicle travel, freight travel, demographics, and street network design. We follow the air pollution literature's standards by using census-tract level aggregations, as fully disaggregate pollution exposure data do not exist. Tracts offer a useful unit of analysis as they roughly represent neighborhoods and generally follow

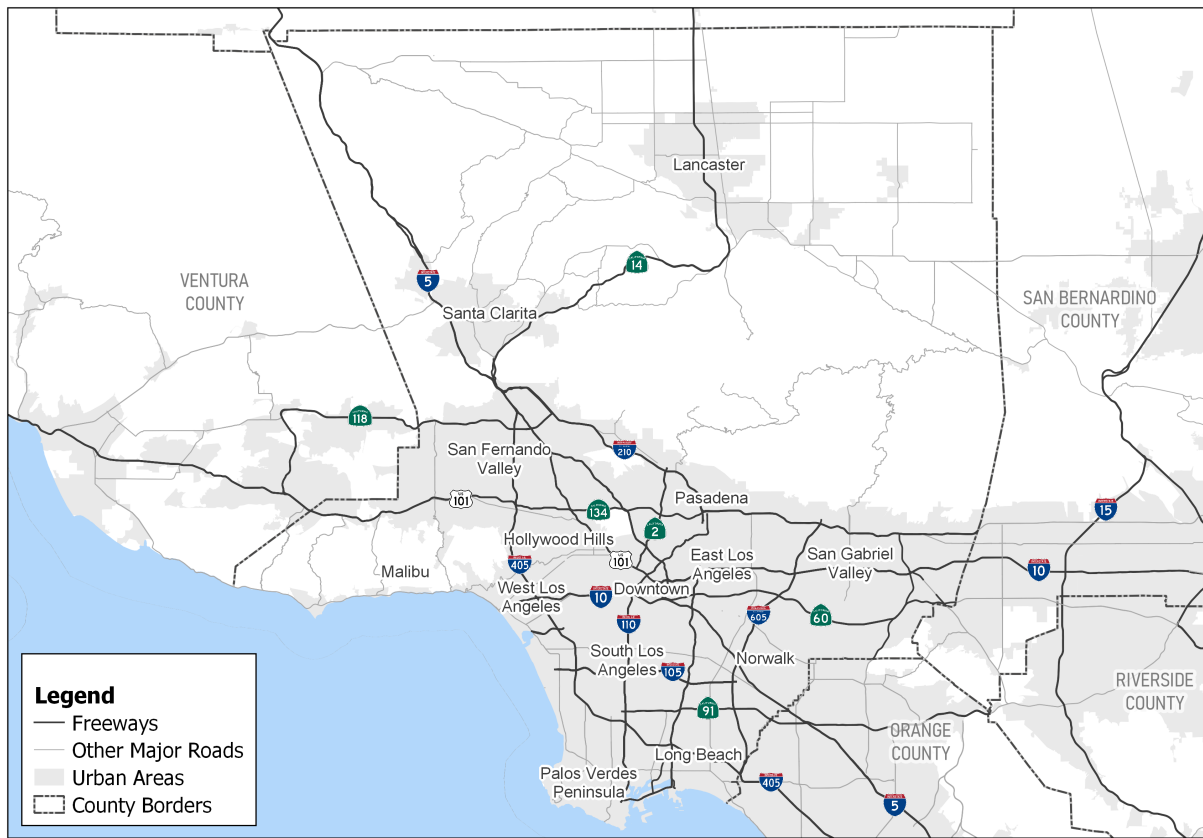


Figure 1. Study Area: Los Angeles County.

real-world physical and social boundaries. Table 1 describes our input data and their units and sources.

3.1.1 Air Pollution

This study uses PM_{2.5} concentrations as the primary air pollution exposure indicator. While CalEnviroScreen and similar national data sets provide PM_{2.5} at sufficiently detailed resolutions, none of them provide emissions from on-road sources only (Diao et al., 2019). However, the Union of Concerned Scientists provides tract-level estimates of PM_{2.5} concentrations from on-road sources in California⁵. The data set combines emissions from on-road sources from the US EPA Emissions Inventory with the InMAP air pollution generation and transport model to estimate tract-level concentrations in California (Reichmuth, 2019b). This allows us to analyze estimated local exposure to vehicular PM_{2.5}.

As a robustness check, we separately use CO₂ emissions from on-road sources, in place of PM_{2.5} concentrations, from the Database of Road Transport Emissions (Hutyra et al., 2015). While the PM_{2.5} data measures concentration as the “stock” of air pollution, the CO₂ data measures tailpipe emissions as the “flow” of vehicular CO₂ into the air at a 1-kilometer resolution. While the PM_{2.5} data use an air transport model of spatial diffusion, the CO₂ data represent emissions from each location without diffusion. Examining this alternative “flow” indicator lets us check the robustness of the results from our primary “stock” indicator.

3.1.2 Passenger Travel

This study uses tract-level resident VKT generation, as a proxy for vehicular pollution production, from the Local Area Transportation Characteristics for Households (LATCH)⁶. This data set combines National Household Travel Survey (NHTS) responses with demographic data from the American Community Survey (ACS) to model household travel behavior and provide tract-level aggregations. This data set is the most comprehensive source of US tract-level VKT estimates. However, LATCH does omit some census tracts such as those comprising military installations or prisons and those without nighttime populations.

We also simulate commuting trips to estimate traffic flows, by race, through different tracts. This uses the LEHD Origin-Destination Employment Statistics (LODES) data set from the Census Bureau⁷. LODES is an administrative enumeration that includes employee home and work census blocks, but only includes private sector employment (Boeing, 2018). This provides us a rough

⁵Union of Concerned Scientists Inequitable Exposure to Air Pollution from Vehicles in California main data (see <https://www.ucsusa.org/resources/inequitable-exposure-air-pollution-vehicles-california-2019>)

⁶US Bureau of Transportation Statistics Local Area Transportation Characteristics for Households (see <https://www.bts.gov/statistical-products/surveys/local-area-transportation-characteristics-households-latch-survey>)

⁷California LEHD LODES7 2017 origin-destination main data (see https://lehd.ces.census.gov/data/lodes/LODES7/ca/od/ca_od_main_JT00_2010.csv.gz)

Table 1. Descriptions, units, and sources of variables used.

Variable	Description	Units	Source
PM _{2.5}	The average concentration of PM _{2.5} in a census tract	micrograms per cubic meter	Union of Concerned Scientists
CO ₂ Emissions	Annual Road transportation emission of CO ₂	metric tons	NASA
Resident VKT generated	The average daily number of kilometers a typical household from the census tract drives	kilometers	LATCH
Proportion White	The share of people in the census tract that are White	n/a	ACS
Median HH Income	The annual median income for households in a census tract	inflation-adjusted 2018 USD, 10,000s	ACS
Truck Traffic Volume	The total number of kilometers that commercial trucks drive through each census tract per day	kilometers	SCAG
Distance to highway	The number of kilometers from the centroid of a tract to the nearest highway	kilometers	Census Tigerline
Intersection Density	Count of intersections per square kilometer	intersections	OpenStreetMap
Grade Mean	Average street segment incline (rise over run)	n/a	OpenStreetMap
Proportion single-family	Share of structures that are single unit detached homes	n/a	ACS
Median Rooms per Home	The median number of rooms per dwelling in a census tract	rooms	ACS
Mean HH Size	The average number of people living in one dwelling per census tract	people	ACS
Population Density	The number of people per square kilometer per census tract	people/kilometer, 1,000s	ACS
Median Home Value	The median value of owner-occupied housing units	inflation-adjusted 2018 USD, 10,000s	ACS

approximation, with some bias, of commute origins and destinations. We aggregate these origins and destinations to the tract-level to match the rest of the input data.

3.1.3 Freight Travel

As discussed in the background, freight traffic can confound analyzing the disparities in the production of and exposure to air pollution. Trucking produces substantial on-road $PM_{2.5}$ which must be controlled for. This study uses estimates from the Southern California Association of Governments (SCAG)⁸. The SCAG model uses employment and population concentrations to estimate truck traffic volume and flow across street network segments in southern California, from which tract-level estimates are generated (SCAG 2010).

3.1.4 Street Network

To control for differences in local street network structure, we measure tract-level intersection density and street grade using OpenStreetMap data and the OSMnx package (Boeing 2017).

3.1.5 Demographics

We collect tract-level race and income data from the 2018 5-year ACS, alongside a set of additional control variables summarized in Table 1. Some of our tracts lack observations across some of these variables. For consistent analysis across multiple specifications, we retain only those tracts with observations across the variables in Table 1.

3.2 Modeling and Analysis

We develop a tract-level indicator, the burden ratio (BR), that links the production of and exposure to vehicular pollution. BR is the ratio of exposure (to vehicular pollution) to production (driving), quantifying tract t 's burden borne in relation to its contribution, as defined in Equation 1:

$$BR_t = \frac{t}{\tau_t} \quad (1)$$

where t is its $PM_{2.5}$ concentration and τ_t is its VKT generated.

3.2.1 OLS Models

We estimate four models: two via OLS and two via GWR. Consider a simple model of the form:

$$\log BR = \beta_0 + \beta X + \epsilon \quad (2)$$

⁸Southern California Association of Governments main data (for more information see <http://www.scag.ca.gov/DataAndTools/Lists/Transportation%20Models%20Tabs/DispForm.aspx?ID=5>)

where $\log \text{BR}$ is the response, β_0 is the intercept, β_1 is a vector of coefficients to be estimated, X is a matrix of n observations on k predictors, and ϵ is the error term. Such a model would predict BR, log-transformed due to its ratio nature, as a function of several predictors. However, given 1) our interest in unpacking the potentially heterogeneous relationship between local $\text{PM}_{2.5}$ concentration and VKT generated, 2) the definition of BR in Equation 1 and 3) the fact that $\log \frac{\text{BR}}{\tau} = \log \text{BR} - \log \tau$, we adapt Equation 2 to arrive at the base model shown in Equation 3:

$$\log \text{BR} = \beta_0 + \beta_1 \log \tau + X + \epsilon \quad (3)$$

Equation 3 is similar to Equation 2 but allows us to estimate coefficient β_1 . We estimate two models via OLS of this general form. In Model 1, X includes a limited set of controls including the White proportion of the population, median household income, resident VKT generation, truck traffic volume, and distance to the nearest highway. In Model 2, X includes a complete set of controls including all those from Model 1 as well as intersection density, mean street grade, proportion of single family homes, median number of rooms per home, mean household size, population density, and median home value.

We log-transform⁹ predictors as needed for a linear relationship and estimate all models with robust standard errors, due to heteroskedasticity. Table 2 presents descriptive statistics of all the variables in their final form.

3.2.2 GWR Models

OLS estimates a global regression model with unvarying parameters across the study region, and is ubiquitous in the environmental justice literature. However, such models cannot unpack spatial variation in statistical relationships. The literature suggests that air pollution exhibits spatial heterogeneity, but sophisticated models to investigate this remain rare in the literature. To unpack any such local variation in coefficients and goodness-of-fit we estimate two GWR models of the general form:

$$\log \text{BR}_t = \beta_{0t} + \beta_{1t} \log \tau_t + X_t + \epsilon_t \quad (4)$$

where $\log \text{BR}_t$ is the response (logarithm of $\text{PM}_{2.5}$ concentration in tract t), β_{0t} is the intercept, β_{1t} is a local coefficient to be estimated, $\log \tau_t$ is the VKT generated by residents of t , X_t is a vector of local coefficients to be estimated, X is a matrix of n observations on k predictors in the local neighborhood of t , and ϵ_t is the local error term. In Model 3, X includes the limited set of controls from Model 1, and in Model 4, X includes the complete set of controls from Model 2.

The GWR models are specified identically to the OLS models, with the same response and predictors. However, their parameter estimates vary by tract location because GWR analysis estimates separate models for the local neighborhood of each tract t in the study region. For each such local regression, observations are weighted by a distance-decay function centered on t . In any application of GWR, the user may choose a fixed bandwidth that is used for every observation, or

⁹An offset of 0.00000001 was added to the variable's value to retain observations with a value of zero in log-transformation.

a variable bandwidth that expands in areas of sparse observations and shrinks in areas of dense observations. Previous studies indicate that choosing a fixed bandwidth is difficult when census enumeration units are used as observations, because these vary in size and shape according to the population density of the area (Mennis and Jordan, 2005; Mennis, 2006). Our study uses a spatially adaptive kernel that adapts for the density of data at each regression location, due to the variability in the sizes of census tracts across Los Angeles County (smaller tracts in urban areas and larger tracts in rural areas). Instead of using a fixed distance, the adaptive kernel determines a fixed number of nearest neighbors to adjust the bandwidth distance accordingly: tracts in dense areas get a narrower bandwidth distance and tracts in sparse areas get a wider one.

The optimal number of nearest neighbors is determined according to the Akaike Information Criterion with small-sample correction (AICc) through an iterative optimization process. A Gaussian function was selected to calculate the weight between tracts with a distance-decay weight closer tracts more than further ones. This resulted in the utilization of 53 and 73 nearest neighbors (tracts) for each local regression in Models 3 and 4, respectively.

3.2.3 Commute Simulation

To examine demographic differences in driving patterns, we use block-level home and employment locations from LODES as a proxy to simulate commuters' routes to work. We solve routes by minimizing (free-flow) travel time and identify, at the home tract level, all the other tracts its trips pass through. We assign each trip as White or non-White probabilistically given the home tract's White population proportion, and adjust the trip counts by the share of commuters who drive to work from each tract. Finally, we sum the total simulated kilometers driven through each tract, by the commuter's race.

To investigate the trends in the disparity with commuting length for each census tract, we developed a subgroup inequity index:

$$F_i = Z_i - T_i \quad (5)$$

where F_i quantifies the degree to which commuting length driven through a census tract is disproportionately made by a specific subgroup i . Here, i can be White and non-White. Z_i is the fraction of the total commuting length made by the subgroup i in this census tract (defined later). T_i is the fraction of commuters who live in this census tract and drive to their workplace who are subgroup i . The sign of this index indicates the direction of the disparity. Positive and negative index values indicate disproportionately high and low representation of commuting length driven through a census tract by subgroup i , respectively.

The fraction of a specific subgroup's commuting length through a census tract is:

$$Z_i = \frac{L_i}{L_T} \quad (6)$$

L_i is the commuting length driven through a certain census tract by subgroup i . L_T is the total commuting length driven through this census tract.

4 Results

4.1 Descriptive Statistics

Table 2 displays the descriptive statistics for response and predictor variables. Figure 2 shows the spatial distribution of BR, PM_{2.5} concentration, household VKT and several sociodemographic indicators across Los Angeles County, in quintiles. Tracts with the greatest BR locate primarily in the region's center, particularly around downtown and South Los Angeles where poorer and non-White residents also concentrate. The correlations between BR and median household income is -0.48 and between BR and proportion White is -0.26. Both are significant indicating a negative relationship between BR and socioeconomic advantage. The tracts in the bottom BR quintile have almost twice the proportion White as the tracts in the top quintile (0.43 vs. 0.23) and have median incomes that are over twice as high (USD 89,100 vs. 45,200). Figure 2 also shows BR's constituent components. PM_{2.5} concentration and household VKT are negatively correlated (-0.49) significant. We unpack these patterns with OLS and GWR regression analysis.

Table 2. Descriptive Statistics for Variables Analyzed, variable unit can be found in Table 1

	Mean	Minimum	Maximum	Std Dev
PM _{2.5}	2.76	0.13	5.48	1.13
Resident VKT generated	60.07	21.95	99.44	11.10
Prop White	0.27	0.00	0.92	0.26
Median HH income	6.68	1.16	22.16	3.19
Truck traffic volume	7806	0	334718	18302
Distance to highway	1.90	0.00	36.33	2.24
Intersection density	49.24	0.05	234.30	20.63
Grade mean	0.02	<0.01	0.10	0.02
Prop single-family	0.54	0.00	1.00	0.29
Median rooms per home	4.74	1.30	10.00	1.12
Mean HH size	3.17	1.36	6.29	0.76
Population density	5.02	0.00	37.83	3.92
Median home value	54.80	1.04	200.00	31.15

4.2 OLS Results

Table 3 reports the OLS regression results for Models 1 and 2. This reveals a significant negative relationship between a tract's vehicular PM_{2.5} concentration and its residents' VKT generation. Controlling for race, income, truck traffic, and highway proximity in Model 1, a 1% increase in VKT generation is associated with a 1.24% decrease in local PM_{2.5} exposure. The full set of controls in Model 2 moderates this effect size somewhat, but a 1% increase in VKT is still associated with a 0.62% decrease in local PM_{2.5} exposure. All else equal, tracts that generate more vehicular travel

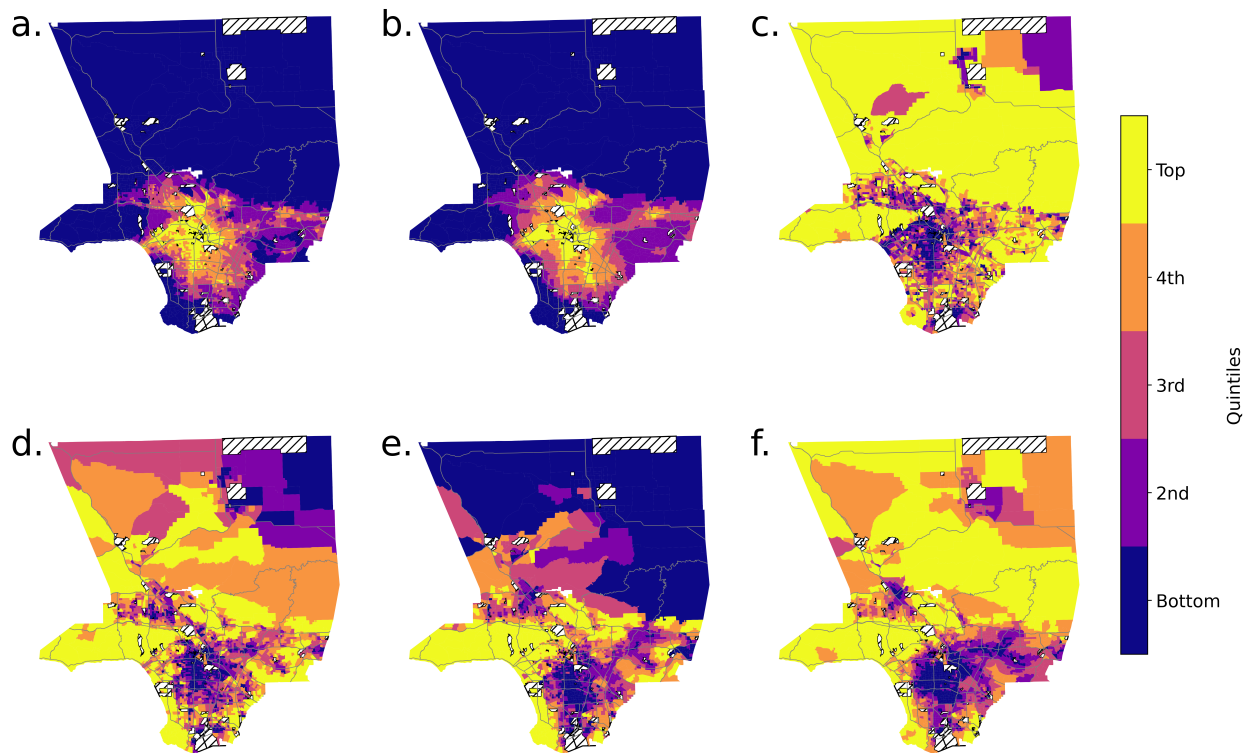


Figure 2. Spatial distribution of tract-level variables across Los Angeles County: a) burden ratio, b) $PM_{2.5}$ Concentration, c) Household VKT, d) Median Household Income, e) Median Home Value, f) Proportion Non-Hispanic White.

tend to be exposed to less vehicular air pollution—an important paradox we unpack in the discussion section.

Further, both Models 1 and 2 reveal a significant negative relationship between a tract's vehicular $PM_{2.5}$ concentration and the White proportion of the population. Even when controlling for income, home value, local traffic, and other covariates, whiter tracts tend to be exposed to less vehicular air pollution. In other words, non-White communities, whether high-income or low-income, are exposed to more $PM_{2.5}$ than otherwise-similar White communities. These models demonstrate that vehicular air pollution burdens distribute inequitably with respect to race across Los Angeles County as a whole.

Our robustness check using CO_2 as the response tells a similar story. A 1% increase in VKT generation is associated with a 0.52% decrease in local CO_2 exposure, or a 0.38% when including the full set of controls. A significant negative relationship is similarly found between the tracts' vehicular CO_2 concentration and the White proportion of the population.

Table 3. Regression model parameter estimates for Model 1 (basic OLS), Model 2 (OLS with full set of controls), Model 3 (basic GWR), and Model 4 (GWR with full set of controls). Standard error is shown in parentheses. Significance denoted as * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

	(1)		(2)		(3)			(4)						
	estimate		estimate		mean	min	max	t<-1.96	t>1.96	mean	min	max	t<-1.96	t>1.96
Intercept	6.13*** (0.29)		1.31*** (0.38)		2.02	-1.58	11.00	2.32%	70.60%	1.37	-1.98	4.84	2.32%	57.42%
log(VKT)	-1.24*** (0.08)		-0.62*** (0.10)		-0.23	-2.79	0.68	38.52%	8.40%	-0.18	-1.46	0.54	33.91%	6.52%
Prop White	-0.61*** (0.06)		-0.63*** (0.07)		-0.14	-1.36	0.76	32.17%	20.38%	-0.03	-0.95	1.22	27.84%	24.98%
Median HH income	0.04*** (0.01)		0.01 (0.01)		-0.01	-0.09	0.25	38.87%	11.66%	0.00	-0.06	0.10	8.76%	7.95%
log(truck traffic volume)	-0.01** (0.01)		0.01* (0.00)		0.00	-0.05	0.05	6.88%	13.94%	0.01	-0.05	0.07	7.86%	21.94%
Distance to highway	-0.12*** (0.01)		-0.09*** (0.01)		-0.05	-0.15	0.02	81.14%	0.09%	-0.04	-0.12	0.02	76.68%	0.09%
Intersection density			-0.00 (0.00)							0.00	-0.01	0.00	11.21%	11.17%
log(grade mean)			-0.06*** (0.01)							-0.04	-0.39	0.14	37.85%	19.71%
Prop single-family			0.44*** (0.07)							0.07	-0.96	0.83	3.66%	19.84%
Median rooms per home			-0.22*** (0.02)							-0.05	-0.44	0.12	43.21%	7.19%
Mean HH size			0.09*** (0.02)							0.04	-0.20	0.33	10.90%	38.20%
log(population density)			0.07*** (0.02)							0.02	-0.10	0.19	1.30%	19.57%
log(median home value)			0.66*** (0.03)							0.05	-0.41	1.19	15.19%	23.82%
N	2238		2238		2238					2238				
R-squared	0.38		0.54		0.67	0.00	0.94			0.74	0.00	0.97		

4.3 GWR Results

Table 3 reports the GWR results for Models 3 and 4 by summarizing the distributions of their coefficients and goodness-of-fit measures. As expected, the GWR models perform substantially better than the OLS models, due to spatial heterogeneity. While Model 1 explains only 38% of the response's variance, Model 3 (with the same predictors) explains 67% of its local variance on average—a substantial improvement. Similarly, Model 2 explains 54% of the response's variance, but Model 4 explains 74% of its local variance on average.

The GWR models unpack the spatial variation of individual predictors across our study area. For example, in Model 1, the White proportion of the population is (uniformly) significantly and negatively associated with vehicular $PM_{2.5}$ exposure. However, Model 3 reveals a more nuanced and spatially varying relationship. Its coefficients for the White proportion of the population range from -1.36 to 0.76 with a mean value of -0.14 (Table 3). The relationship between these variables is not stationary across our study region, though it tends to be negative on average—similar to what Models 1 and 2 suggest globally. These variations in both significance and direction highlight the importance of assessing these relationships locally as well as globally.

Figure 3 depicts Model 4's spatial distributions of local t -statistics and R^2 values across the study region. The combined statistical relationship of the predictors with the response varies across the region. The figure illustrates where exactly the GWR models offer a better fit than the OLS models could. Overall, the GWR R^2 values improve on the OLS values in most tracts.

The consistency of these model results across specifications lends confidence in their estimates. In Model 3, 39% of tracts show a significant, negative relationship between resident VKT generation and vehicular $PM_{2.5}$ exposure, while only 8% of tracts show significant, positive relationship. With the full set of controls in Model 4, 34% of tracts show a significant, negative relationship between resident VKT generation and vehicular $PM_{2.5}$ exposure, while only 7% of tracts show significant, positive relationship. Figure 3 illustrates where these relationships tend to be negative, including the low-income, Latino eastside of Los Angeles and the high-income, White communities in the Hollywood Hills and along the oceanfront. Such areas demonstrate the greatest injustice as the negative relationship means that those who drive the most are exposed to the least pollution, and vice versa.

4.4 Commute Simulation Results

The commute simulation results reveal unequal patterns in driving between White and non-White commuters, which helps explain the exposure disparity. Overall figures do not suggest a particularly large inequality: While the county's population is 26% White, and 30% of all commuter kilometers driven are attributable to White people, White commuters drive through tracts that are on average 61% non-White to get to work, but non-White commuters traverse tracts that are only 26% White. Countywide, the population-weighted value of the index is 0.0153, implying that most tracts in the county experienced disproportionately many trips from White commuters relative to local population. This disparity is greater in census tracts with highways, reaching 0.031 in tracts with highways but only 0.0147 in tracts without highways.

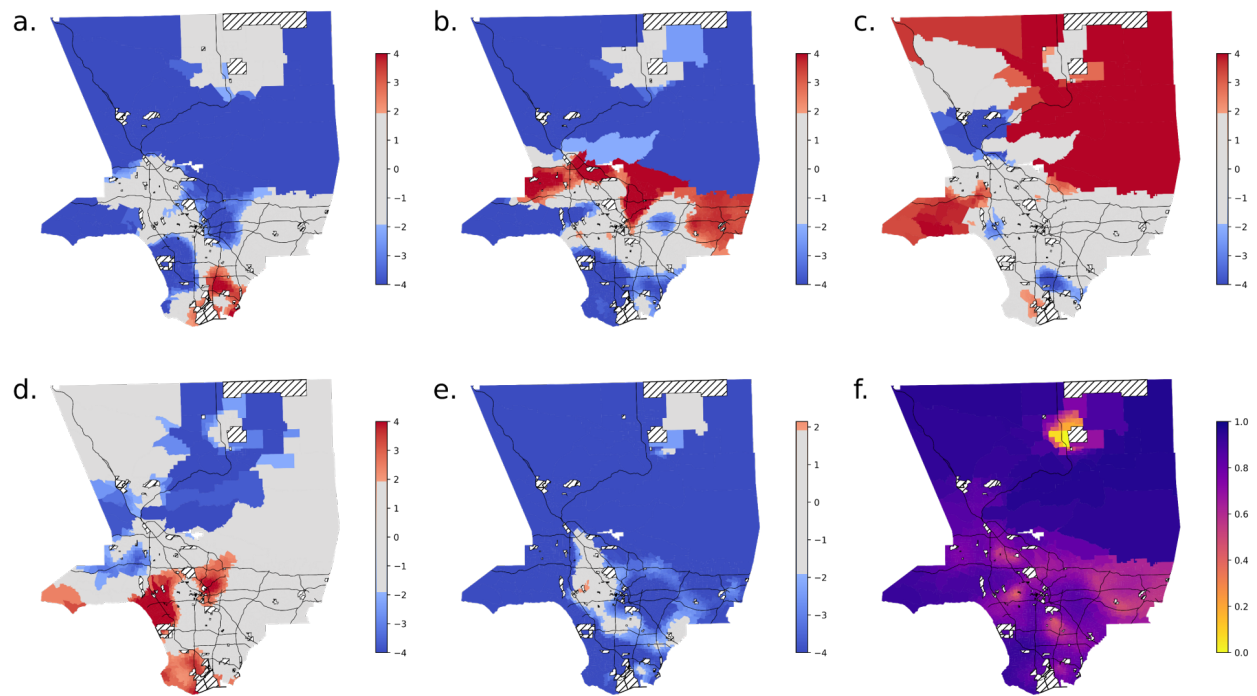


Figure 3. Distribution of local t-values for parameter estimates and local R-squared by census tract for complete GWR model: a. log(VKT); b. Proportion White; c. Median HH income; d. log(truck traffic volume); e. Distance to highway; f. local R-squared.

Figure 4 shows the spatial distribution of the subgroup inequity index. Positive values mean the local commuters living in a census tract are less White than the commuters driving through this tract, while negative values mean the opposite. Figure 4 unpacks the disparity and illustrates the role of historical highway placement in racialized driving patterns. Values tend to be positive in low-income, Latino and Black communities in South Central Los Angeles, eastside Los Angeles, San Gabriel Valley, and San Fernando Valley, especially along the major highways. Negative values appear in high-income, White communities in the Hollywood Hills and along the oceanfront. Major highways stand out in particular: Tracts within a 1/8-mile of Interstate 10 - the region's east-west thoroughfare - have a population-weighted value of 0.096, while those near Interstates 110 and 105 - the largest freeways crossing the predominately Black subregion of South Los Angeles - have a value of 0.129. These discrepancies - implying that the population driving through those tracts is ten or more percentage points Whiter than the resident population - are remarkable in light of the countywide population being less than one third White.

At the same time, there are examples for major freeways on which commuters resemble or are less white than the resident populations of their surrounding areas: For Interstate 405 and US 101, our inequity index takes values of -0.001 and -0.099, respectively. Both share in common that they run through mountain passes in the majority-white Hollywood Hills - the Sepulveda and Cahuenga passes - where alternative alignments would almost certainly have been inconvenient or cost-prohibitive.

Figure 5 illustrates a typical example of the commute routes from a high-income, White tract contrasted with those from a low-income, non-White tract. White commutes tend to traverse less-White tracts in downtown, South Los Angeles, and the San Fernando Valley. However, non-White commutes rarely traverse heavily-White peripheral neighborhoods, such as those in the Hollywood Hills and oceanfront communities.

5 Discussion

The research literature has explored race and class disparities in local vehicular air pollution exposure, but without clearly accounting for the local residents' own contribution to their exposure. In this study, we ask: are different communities exposed to such pollution at a level proportional to how much they drive? We find that—all else equal—tracts whose residents drive less experience more air pollution. Furthermore, tracts with a larger non-White population share—whether high income or low income—experience more air pollution than do Whiter but otherwise similar tracts. Comparable trends hold in the GWR models on average as well. This reveals an injustice in pollution burden with a distinct racial dimension, and these results are robust across multiple model specifications and both global and local estimation techniques.

However, some limitations in our study open the door for future research. While the simulation analysis provides evidence that unequal patterns in driving could be generating some of the observed disparity in exposure to and production of pollution, it cannot rule out the influence of other mechanisms such as differences in wind patterns (e.g., [Heblich et al., 2021](#)) or differences in fuel efficiency across tracts as potentially important mechanisms.

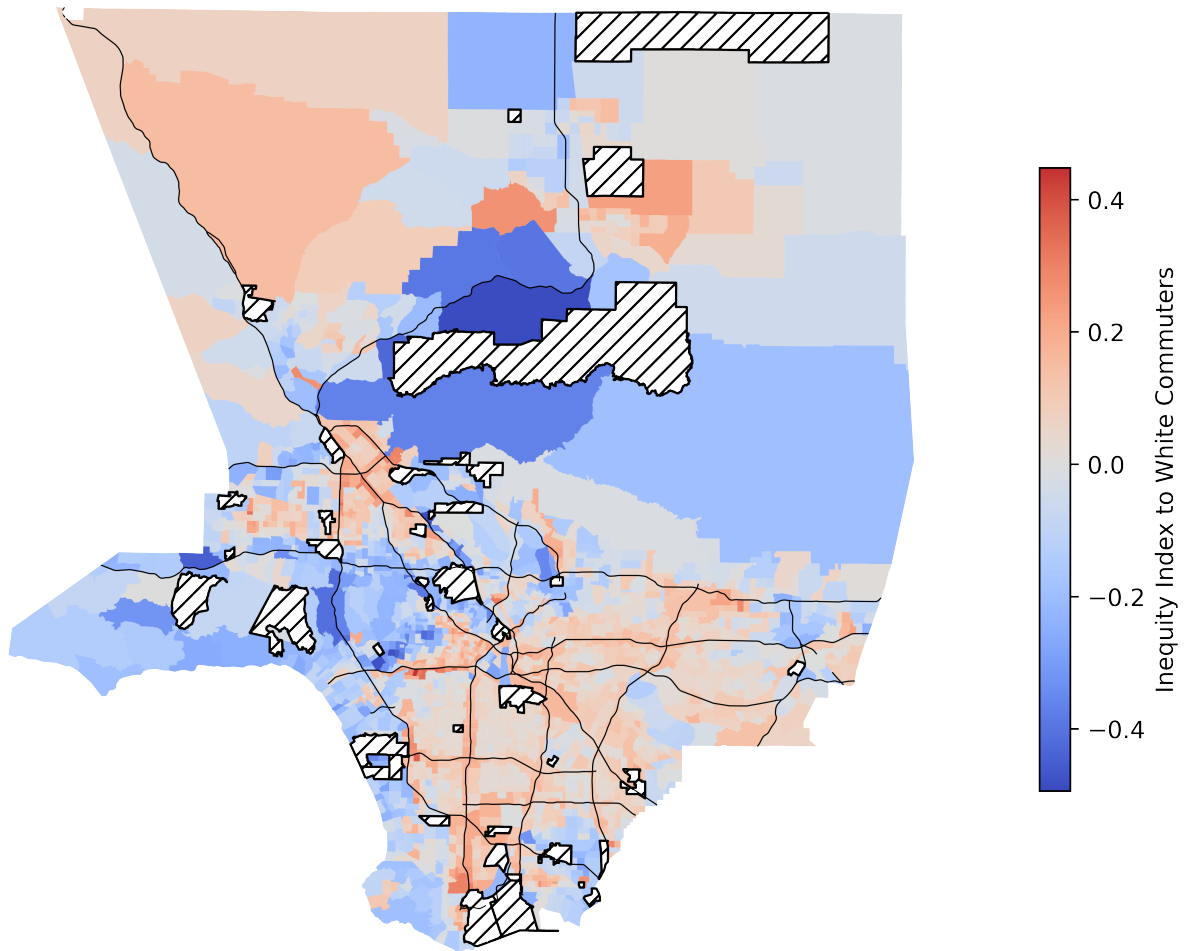


Figure 4. Distribution of local inequity index to White commuters and its relative location to highways.

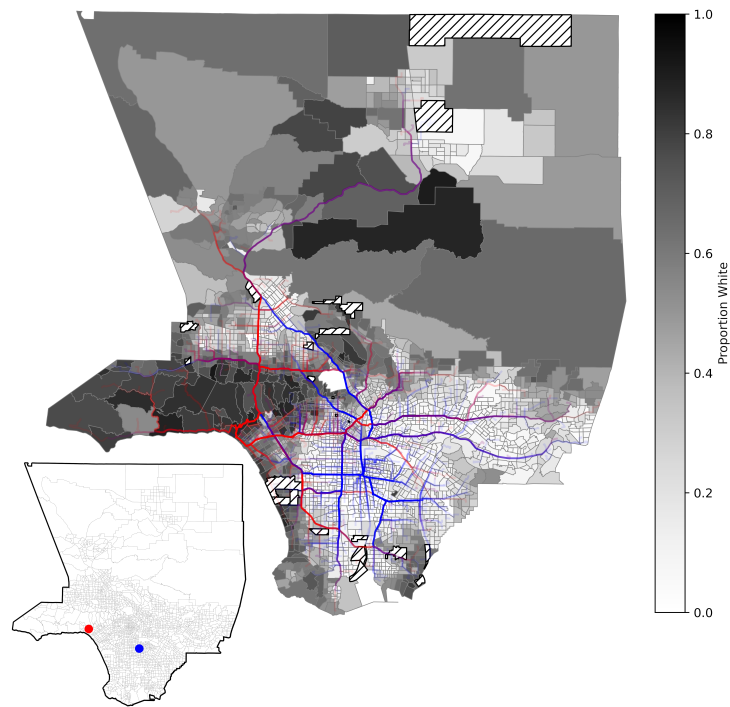


Figure 5. Commute routes for a typical high-income White tract and low-income minority tract: red represents routes from a majority White tract, blue represents routes from a majority non-White tract, and the grey tract colors represent the proportion White in the tract. The inset map shows the location of the two case study tracts.

Nevertheless, our models consistently reveal the same broad story: tracts with a greater share of White residents and tracts with residents who drive more are exposed to less vehicular air pollution. One natural way to interpret this relationship is that people from majority-White neighborhoods do most of their “extra” driving through minority neighborhoods so that they are less exposed to pollution at home despite producing more of it. However, there are other potential mechanisms that could explain the relationship. For example, majority-White neighborhoods could be located in areas where the prevailing winds push the pollution away from their neighborhood and toward minority neighborhoods, or people from majority-White neighborhoods could drive more fuel-efficient cars so that even though they drive more, they produce and are exposed to less pollution at home. However, the history of highway construction through minority neighborhoods strongly suggests that differences in where car travel occurs is driving the regression results.

The commute simulation provides additional evidence on the source of the pollution disparity, and illustrates the role that highways play in mediating the disparity. Its results support the theory that disparities in exposure to pollution result from where people drive. On average, when White people drive to work they traverse tracts that are far more non-White than the tracts where most White people live. They do so predominantly through tracts that contain highways. In other words, White commuters receive the benefits of driving on a highway, but because those highways are predominantly in non-White neighborhoods, other racial groups bear some of the costs of that driving.

This disparity is less common in the opposite direction: on average, non-White commuters do not travel through tracts that are substantially Whiter than their home tracts - and where they do, it is on freeway segments that follow alignments made necessary by local topography such as with mountain passes. By not building the Beverly Hills Freeway, a substantial share of commuters who would have taken this route likely drive through majority non-white Mid-City on Interstate 10 instead (Masters, 2014); a similar redirection of Interstate 405 or US 101 would not be possible without far greater and costlier detours.

Overall, our findings reveal a systematic environmental injustice: transportation infrastructure (i.e., highways) benefits wealthy populations who are often White while extracting a cost on minorities and lower-income populations. A large body of existing research had documented the negative health and developmental impacts of exposure to air pollution. In Los Angeles, minority and low income residents are disproportionately burdened by this air pollution and the health impacts that accompany it: minority communities in Los Angeles experience more vehicular pollution than their White neighbors, while driving less. Furthermore, we provide evidence that this disparity is at least partially driven by the fact that commuters from heavily White communities drive disproportionately through non-White areas to get to work. This commuting pattern is influenced by the historical placement of highways since when White people commute through minority communities, they largely do so on highways. In other words, past planning decisions allow richer and Whiter residents to pollute minority communities as they commute to jobs in the urban core, while living in less polluted peripheral areas.

As past policymakers bear some responsibility for this environmental injustice, current policymakers have a duty to address the issue. A particular challenge in addressing this injustice is that infrastructure is not easy to move and rarely are highways dismantled after they have been built. The

most effective options will likely be policies that either reduce the use of highways and vehicles or reduce the impact of vehicle usage. One option state policymakers could take would be to continue raising the fuel efficiency standard for new cars and encouraging the electrification of vehicles. Being aggressive in environment protection, California has issued an executive order requiring sales of all new passenger vehicles to be zero-emission by 2035¹⁰. Both of these approaches would reduce the amount of pollution emitted from a vehicle without necessarily reducing the amount of driving. These plans would initially benefit richer and Whiter communities as they will likely be the first to purchase these new cars, but the reduced emissions would also benefit minority communities when those commuters drive through these areas. However, this will not address the pollution from tire wear and brake dust, and if power stations are located in minority communities and they emit their own air pollution, then electrification could just be trading one source of pollution for another.

Policymakers could also enact tolls or other forms of congestion taxes to reduce total driving (Schindler and Caruso, 2021) or capture its externalities. Cities such as London and New York have implemented or are in the process of implementing congestion tolls for drivers entering their downtown's and Los Angeles could adopt a similar plan. Research on the London congestion toll has found that it reduced traffic within central London, increased traffic speeds, and increased bus ridership (Leape, 2006). If the taxes or tolls apply to everyone driving in the downtown area, it could be regressive as many minority and low income residents live and drive in that area. However, it could be more progressive if it only applies to people entering the downtown area from more affluent areas. Such tolls could reduce car commuting and encourage public transit alternatives, and the revenue from the tax could be used to further improve public transit provision.

Another potential option could be discouraging commuting altogether by incentivizing more people to work from home. The COVID-19 pandemic saw a surge in workers working from home and continuing this trend after the pandemic could reduce the amount of commuting through minority neighborhoods and reduce air pollution. City or state policymakers could offer tax credit to encourage this. However, people who are able to work from home are primarily Whiter and more affluent, so such tax credits could be a regressive method to encourage working from home. There is also little evidence of the effectiveness of incentives to encourage working remotely.

Policymakers could also address the environmental injustice through the housing market. Proposed state legislation would permit more construction in job rich areas, which could reduce the amount of driving needed to get to work¹¹. Legalizing the construction of denser and more affordable housing in less polluted areas could also reduce exposure disparities by increasing access to those areas for low income people. Improving access to healthy areas could benefit the people who are able to move, but it could increase commuting lengths and potentially exacerbate pollution exposure for those left behind in more heavily polluted neighborhoods.

¹⁰Governor Newsom's Zero-Emission by 2035 Executive Order (N-79-20) (see <https://ww2.arb.ca.gov/resources/fact-sheets/governor-newsoms-zero-emission-2035-executive-order-n-79-20>)

¹¹For more details on state legislation see, https://leginfo.legislature.ca.gov/faces/billAnalysisClient.xhtml?bill_id=201920200SB50

6 Conclusion

This paper extends quantitative research on transportation-environmental justice through a case study of the production of and exposure to vehicular pollution in Los Angeles. We assess local exposure (to vehicular pollution) adjusted by local production (VKT) through four models using two estimation methods: OLS and GWR. OLS demonstrates the global relationships between response variable and predictor variables. GWR allows for a localized regression models that better capture spatial heterogeneity and explain the response's variance. The visualization of the local parameter estimates, t test results, and model R^2 values from the GWR help us investigate the effect of various explanatory factors on the magnitude of environmental justice and identify their spatial patterns.

We found that tracts that generate more vehicular travel tend to be exposed to less vehicular air pollution, all else equal. Race and income significantly predict pollution exposure, even when controlling for a full set of related variables. The commute simulation helps explain these findings by illustrating the role of freeways in commuting patterns. On average, White commuters traverse tracts that are far more non-White than the tracts where they live, but non-White commuters do not travel through tracts that are substantially Whiter than their own. In concert, these findings extend the transportation-environmental justice literature with more nuanced geographical analysis methods. As the injustices found here largely result from historical policy and planning, it is up to today's policymakers and planners to understand and attenuate them.

References

- Alman, B. L., Stingone, J. A., Yazdy, M., Botto, L. D., Desrosiers, T. A., Pruitt, S., Herring, A. H., Langlois, P. H., Nembhard, W. N., Shaw, G. M., Olshan, A. F., and Luben, T. J. (2019). Associations between PM_{2.5} and risk of preterm birth among liveborn infants. *Annals of Epidemiology*, 39:46–53.e2.
- Anderson, J. O., Thundiyil, J. G., and Stolbach, A. (2012). Clearing the Air: A Review of the Effects of Particulate Matter Air Pollution on Human Health. *Journal of Medical Toxicology*, 8(2):166–175.
- Bae, C.-H. C., Sandlin, G., Bassok, A., and Kim, S. (2007). The exposure of disadvantaged populations in freeway air-pollution sheds: A case study of the seattle and portland regions. *Environment and Planning B: Planning and Design*, 34(1):154–170.
- Baum-Snow, N. (2007). Did Highways Cause Suburbanization? *The Quarterly Journal of Economics*, 122(2):775–805.
- Beelen, R., Hoek, G., van den Brandt, P. A., Goldbohm, R. A., Fischer, P., Schouten, L. J., Armstrong, B., and Brunekreef, B. (2008). Long-term exposure to traffic-related air pollution and lung cancer risk. *Epidemiology*, 19(5):702–710.
- Bell, M. L. and Ebisu, K. (2012). Environmental Inequality in Exposures to Airborne Particulate Matter Components in the United States. *Environmental Health Perspectives*, 120(12):1699–1704.
- Beris, A. N. and Edwards, B. J. (1994). *Thermodynamics of flowing systems: with internal microstructure*. Number 36 in Oxford engineering science series. Oxford University Press, New York.

- Board, C. A. R. (2020). Current California GHG Emission Inventory Data | California Air Resources Board.
- Boeing, G. (2017). Osmnx: New methods for acquiring, constructing, analyzing, and visualizing complex street networks. *Computers, Environment and Urban Systems*, 65:126–139.
- Boeing, G. (2018). Estimating local daytime population density from census and payroll data. *Regional Studies, Regional Science*, 5:179–182.
- Bravo, M. A., Anthopoulos, R., Bell, M. L., and Miranda, M. L. (2016). Racial isolation and exposure to airborne particulate matter and ozone in understudied US populations: Environmental justice applications of downscaled numerical model output. *Environment International*, 92-93:247–255.
- Brugge, D., Durant, J. L., and Rioux, C. (2007). Near-highway pollutants in motor vehicle exhaust: A review of epidemiologic evidence of cardiac and pulmonary health risks. *Environmental Health*, 6(1):23.
- Brunekreef, B., Janssen, N. A. H., de Hartog, J., Harssema, H., Knape, M., and van Vliet, P. (1997). Air pollution from truck traffic and lung function in children living near motorways. *Epidemiology*, 8(3):298–303.
- Button, K. (1990). Environmental externalities and transport policy. *Oxford Review of Economic Policy*, 6(2):61–75.
- Dear, M. (2001). *From Chicago to L.A.: Making Sense of Urban Theory*. Sage Publications, Thousand Oaks, CA.
- Delucchi, M. A. (2000). Environmental externalities of motor-vehicle use in the us. *Journal of Transport Economics and Policy*, 34(2):135–168.
- Diao, M., Holloway, T., Choi, S., O'Neill, S. M., Al-Hamdan, M. Z., Donkelaar, A. V., Martin, R. V., Jin, X., Fiore, A. M., Henze, D. K., Lacey, F., Kinney, P. L., Freedman, F., Larkin, N. K., Zou, Y., Kelly, J. T., and Vaidyanathan, A. (2019). Methods, availability, and applications of PM2.5 exposure estimates derived from ground measurements, satellite, and atmospheric models. *Journal of the Air & Waste Management Association*, 69(12):1391–1414.
- DiMento, J. F. C. (2009). Stent (or dagger?) in the heart of town: Urban freeways in syracuse, 1944—1967. *Journal of Planning History*, 8(2):133–161.
- Estrada, G. (2005). If you build it, they will move: The los angeles freeway system and the displacement of mexican east los angeles, 1944-1972. *Southern California Quarterly*, pages 287–315.
- Gauderman, W. J., Avol, E., Lurmann, F., Kuenzli, N., Gilliland, F., Peters, J., and McConnell, R. (2005). Childhood Asthma and Exposure to Traffic and Nitrogen Dioxide:. *Epidemiology*, 16(6):737–743.
- Giuliano, G. (2003). Travel, location and race/ethnicity. *Transportation Research Part A" Policy and Practice*, 37:351–372.
- Habre, R., Girguis, M., Urman, R., Fruin, S., Lurmann, F., Shafer, M., Gorski, P., Franklin, M., McConnell, R., Avol, E., and Gilliland, F. (2021). Contribution of tailpipe and non-tailpipe traffic sources to quasi-ultrafine, fine and coarse particulate matter in southern California. *Journal of the Air & Waste Management Association*, 71(2):209–230.

- Hasheminassab, S., Daher, N., Ostro, B. D., and Sioutas, C. (2014). Long-term source apportionment of ambient fine particulate matter (PM 2.5) in the Los Angeles Basin: A focus on emissions reduction from vehicular sources. *Environmental Pollution*, 193:54–64.
- Heblich, S., Trew, A., and Zylberberg, Y. (2021). East-side story: Historical pollution and persistent neighborhood sorting. *Journal of Political Economy*, 129(5):1508–1552.
- Houston, D., Krudysz, M., and Winer, A. (2008). Diesel truck traffic in low-income and minority communities adjacent to ports: Environmental justice implications of near-roadway land use conflicts. *Transportation Research Record*, 2067(1):38–46.
- Houston, D., Li, W., and Wu, J. (2014). Disparities in Exposure to Automobile and Truck Traffic and Vehicle Emissions Near the Los Angeles–Long Beach Port Complex. *American Journal of Public Health*, 104(1):156–164.
- Houston, D., Wu, J., Ong, P., and Winer, A. (2004). Structural disparities of urban traffic in southern california: Implications for vehicle-related air pollution exposure in minority and high-poverty neighborhoods. *Journal of Urban Affairs*, 26:565–592.
- Hutyra, L. R., Gately, C., and Wing, I. S. (2015). CMS: DARTE Annual On-road CO2 Emissions on a 1-km Grid, Conterminous USA, 1980-2012. *ORNL DAAC*.
- Jagai, J. S., Messer, L. C., Rappazzo, K. M., Gray, C. L., Grabich, S. C., and Lobdell, D. T. (2017). County-level cumulative environmental quality associated with cancer incidence: Environment and Cancer Incidence. *Cancer*, 123(15):2901–2908.
- Janssen, N. A., van Vliet, P. H., Aarts, F., Harssema, H., and Brunekreef, B. (2001). Assessment of exposure to traffic related air pollution of children attending schools near motorways. *Atmospheric Environment*, 35(22):3875–3884.
- Jorgenson, A. K., Hill, T. D., Clark, B., Thombs, R. P., Ore, P., Balistreri, K. S., and Givens, J. E. (2020). Power, proximity, and physiology: does income inequality and racial composition amplify the impacts of air pollution on life expectancy in the United States? *Environmental Research Letters*, 15(2):024013.
- Karner, A. A., Eisinger, D. S., and Niemeier, D. A. (2010). Near-roadway air quality: synthesizing the findings from real-world data. *Environmental science & technology*, 44(14):5334–5344.
- Künzli, N., McConnell, R., Bates, D., Bastain, T., Hricko, A., Lurmann, F., Avol, E., Gilliland, F., and Peters, J. (2003). Breathless in Los Angeles: The Exhausting Search for Clean Air. *American Journal of Public Health*, 93(9):1494–1499.
- Leape, J. (2006). The london congestion charge. *Journal of Economic Perspectives*, 20(4):157–176.
- Lee, S. L., Wong, W. H. S., and Lau, Y. L. (2006). Association between air pollution and asthma admission among children in Hong Kong. *Clinical & Experimental Allergy*, 36(9):1138–1146.
- Liévanos, R. S. (2019). Racialized Structural Vulnerability: Neighborhood Racial Composition, Concentrated Disadvantage, and Fine Particulate Matter in California. *International Journal of Environmental Research and Public Health*, 16(17):3196.
- Ljubimova, J. Y., Braubach, O., Patil, R., Chiechi, A., Tang, J., Galstyan, A., Shatalova, E. S., Kleinman, M. T., Black, K. L., and Holler, E. (2018). Coarse particulate matter (PM2.5–10) in Los Angeles Basin air induces expression of inflammation and cancer biomarkers in rat brains. *Scientific Reports*, 8(1):5708.

- Loftus, C. T., Ni, Y., Szpiro, A. A., Hazlehurst, M. F., Tylavsky, F. A., Bush, N. R., Sathyanarayana, S., Carroll, K. N., Young, M., Karr, C. J., and LeWinn, K. Z. (2020). Exposure to ambient air pollution and early childhood behavior: A longitudinal cohort study. *Environmental Research*, 183:109075.
- Masters, N. (2014). Why Isn't There a Freeway to Beverly Hills? *KCET*. Section: Food & Discovery.
- Mennis, J. (2006). Mapping the results of geographically weighted regression. *The Cartographic Journal*, 43(2):171–179.
- Mennis, J. L. and Jordan, L. (2005). The distribution of environmental equity: Exploring spatial nonstationarity in multivariate models of air toxic releases. *Annals of the Association of American Geographers*, 95(2):249–268.
- Mikati, I., Benson, A. F., Luben, T. J., Sacks, J. D., and Richmond-Bryant, J. (2018). Disparities in Distribution of Particulate Matter Emission Sources by Race and Poverty Status. *American Journal of Public Health*, 108(4):480–485.
- Mohl, R. A. (2004). Stop the Road: Freeway Revolts in American Cities. *Journal of Urban History*, 30(5):674–706.
- Oosterlee, A., Drijver, M., Lebet, E., and Brunekreef, B. (1996). Chronic respiratory symptoms in children and adults living along streets with high traffic density. *Occupational and environmental medicine*, 53:241–7.
- Pan, H., Bartolome, C., Gutierrez, E., Princevac, M., Edwards, R., Boarnet, M. G., and Houston, D. (2013). Investigation of roadside fine particulate matter concentration surrounding major arterials in five southern californian cities. *Journal of the Air & Waste Management Association*, 63(4):482–498.
- Perez, J. (2017). The Los Angeles Freeway and the History of Community Displacement. *Toro Historical Review*, 3.
- Pershagen, G., Rylander, E., Norberg, S., Eriksson, M., and Nordvall, S. L. (1995). Air Pollution Involving Nitrogen Dioxide Exposure and Wheezing Bronchitis in Children. *International Journal of Epidemiology*, 24(6):1147–1153.
- PolicyLink and USC Program for Environmental and Regional Equity (2017). An Equity Profile of the Los Angeles Region.
- Poorfakhraei, A., Tayarani, M., and Rowangould, G. (2017). Evaluating health outcomes from vehicle emissions exposure in the long range regional transportation planning process. *Journal of Transport & Health*, 6:501–515.
- Pratt, G., Vadali, M., Kvale, D., and Ellickson, K. (2015). Traffic, Air Pollution, Minority and Socio-Economic Status: Addressing Inequities in Exposure and Risk. *International Journal of Environmental Research and Public Health*, 12(5):5355–5372.
- Pulido, L. (2000). Rethinking environmental racism: White privilege and urban development in southern california. *Annals of the Association of American Geographers*, page 30.
- Reichmuth, D. (2019a). Inequitable Exposure to Air Pollution from Vehicles in California. Technical report, Union of Concerned Scientists, Cambridge, MA.
- Reichmuth, D. (2019b). Inequitable exposure to air pollution from vehicles in california fact sheet. Technical report, Union of Concerned Scientists.

- Rowangould, D., Karner, A., and London, J. (2016). Identifying environmental justice communities for transportation analysis. *Transportation Research Part A: Policy and Practice*, 88:151–162.
- Rowangould, G. (2013). A census of the us near-roadway population: Public health and environmental justice considerations. *Transportation Research Part D Transport and Environment*, 25:59–67.
- Rowangould, G. (2015). A new approach for evaluating regional exposure to particulate matter emissions from motor vehicles. *Transportation Research Part D-transport and Environment*, 34:307–317.
- SCAG (2010). Heavy Duty Truck Model. Technical report, Southern California Association of Governments.
- Schindler, M. and Caruso, G. (2021). Urban interventions to reduce pollution exposure and improve spatial equity. *Geographical Analysis*, n/a(n/a).
- Schweitzer, L. and Valenzuela, A. (2004). Environmental Injustice and Transportation: The Claims and the Evidence. *Journal of Planning Literature*, 18(4):383–398.
- Seagram, A. F., Brown, S. G., Huang, S., Landsberg, K., and Eisinger, D. S. (2019). National assessment of near-road air quality in 2016: Multi-year pollutant trends and estimation of near-road pm2.5 increment. *Transportation Research Record*, 2673(2):161–171.
- Sunyer, J., Esnaola, M., Alvarez-Pedrerol, M., Forn, J., Rivas, I., López-Vicente, M., Suades-González, E., Foraster, M., Garcia-Esteban, R., Basagaña, X., and et al. (2015). Association between traffic-related air pollution in schools and cognitive development in primary school children: A prospective cohort study. *PLOS Medicine*, 12(3):e1001792.
- Tessum, C. W., Paoella, D. A., Chambliss, S. E., Apte, J. S., Hill, J. D., and Marshall, J. D. (2021). Pm2.5 pollutants disproportionately and systemically affect people of color in the united states. *Science Advances*, 7(18).
- US EPA, O. (2018). Fast Facts on Transportation Greenhouse Gas Emissions.
- White House (2021). Memorandum on redressing our nation’s and the federal government’s history of discriminatory housing practices and policies. *The White House*.
- World Health Organization, editor (2006). *Air quality guidelines: global update 2005: particulate matter, ozone, nitrogen dioxide, and sulfur dioxide*. World Health Organization, Copenhagen, Denmark. OCLC: ocn162119780.
- Wu, J., Houston, D., Lurmann, F., Ong, P., and Winer, A. (2009). Exposure of pm2.5 and ec from diesel and gasoline vehicles in communities near the ports of los angeles and long beach, california. *Atmospheric Environment*, 43:1962–1971.
- Yin, Y.-W., Cheng, J.-P., Duan, Y.-S., Wei, H.-P., Ji, R.-X., Yu, J.-L., and Yu, H.-R. (2011). [Correlation analysis between the PM2.5, PM10 which were taken in the hazy day and the number of outpatient about breathing sections, breathing sections of pediatrics in Shanghai]. *Huan Jing Ke Xue= Huanjing Kexue*, 32(7):1894–1898.
- Yuan, Q. (2018). Mega freight generators in my backyard: A longitudinal study of environmental justice in warehousing location. *Land Use Policy*, 76:130–143.
- Zhu, Y., Hinds, W. C., Kim, S., and Sioutas, C. (2002). Concentration and size distribution of ultrafine particles near a major highway. *Journal of the Air & Waste Management Association (1995)*, 52(9):1032—1042.

Data Management Plan

Products of Research

No primary data were collected for this study, only secondary data.

Data Format and Content

1. Air Pollution data

1.1. Union of Concerned Scientists (pm_ucs)

DATA STORAGE:

- Data Format: CSV
- Find data at: <https://www.ucsusa.org/resources/inequitable-exposure-air-pollution-vehicles-california-2019>

DATA DESCRIPTION:

- Pre-existing data collected by a third-party
- Modeling and simulation data: Input data and outputs for models such as CUBE, TRANSCad, or VISSIM; input data and output for simulation models created by PSR researchers.
- Contextual documentation: Included in project input folder

METADATA INFORMATION:

- Title: Inequitable Exposure to Air Pollution from Vehicles in California
- Year: 2019
- Author: Reichmuth, David. Union of Concerned Scientists
- Abstract: A quantified the exposure of different groups to particulate matter (PM2.5) from on-road sources and compared it to other demographic segments. Beyond the scope of this analysis, emissions from ports, agricultural practices, dust, and other sources are well known to contribute to poor air quality and negative health outcomes for affected areas.
- Methods: To assess chronic exposure to particulate air pollution in California, the Union of Concerned Scientists used a model of pollutant generation and transport (InMAP) to generate estimates of average annual concentrations of particulate matter smaller than 2.5 micrometers (PM2.5) in California with a maximum resolution of one square kilometer (km²). InMAP models both the transport of primary PM2.5 emissions and the formation of secondary PM2.5 through atmospheric reactions. Tailpipe and refueling emissions (nitrogen oxides, sulfur oxides, PM2.5, and volatile organics) from on-road vehicles were adapted from the US Environmental Protection Agency (EPA) National Emissions Inventory (NEI).
- Link to data source: <https://www.ucsusa.org/resources/inequitable-exposure-air-pollution-vehicles-california-2019>

1.2. DARTE Highway Data (co2_darte)

DATA STORAGE:

- Data Format: Tag Image File Format
- Find data at: https://daac.ornl.gov/cgi-bin/dsviewer.pl?ds_id=1735

DATA DESCRIPTION:

- Pre-existing data collected by a third-party
- Vehicle Miles Traveled Data: Vehicle miles traveled (VMT) per capita is calculated as the total annual miles of vehicle travel divided by the total population in a state or in an urbanized area. Data for this indicator come from the Federal Highway Administration (FHWA), 2011 Highway Statistics. The reports are based on individual state reports on traffic data counts collected through permanent automatic traffic recorders on public roadways. Data on VMT for urbanized areas are available from the FHWA Highway Statistics Series. These data are calculated as the total daily miles of vehicle travel in an urbanized area divided by the total population. An urbanized area is defined as an area with 50,000 persons that at a minimum encompasses the land area delineated as the urbanized area by the U.S. Census Bureau.
- Contextual documentation (like data dictionaries): Included on website

METADATA INFORMATION:

- Title: DARTE Annual On-road CO2 Emissions on a 1-km Grid, Conterminous USA, V2, 1980-2017
- Year: 2019
- Author: Gately, C., L.R. Hutyra, and I.S. Wing, ORNL DAAC
- Abstract: This data set provides a 38-year, 1-km resolution inventory of annual on-road CO2 emissions for the conterminous United States based on roadway-level vehicle traffic data and state-specific emissions factors for multiple vehicle types on urban and rural roads as compiled in the Database of Road Transportation Emissions (DARTE). CO2 emissions from the on-road transportation sector are provided annually for 1980-2017 as a continuous surface at a spatial resolution of 1 km.
- Methods: DARTE combined the Federal Highway Administration's (FHWA's) Highway Performance Monitoring System (HPMS) roadway-level vehicle miles traveled with year- and state-specific emissions factors for five vehicle types on six classes of urban and rural roads. Vehicle emissions were estimated directly at the scale of individual road segments without the need to downsize emissions using spatial predictors. This approach helped to refine the uncertainty in the spatial distribution of road vehicle emissions and showed the highly nonlinear relationship between population density and emissions.
- Link to data source: https://daac.ornl.gov/CMS/guides/CMS_DARTE_V2.html

2. Passenger Travel

2.1. LATCH Data (vmt_latch)

DATA STORAGE:

- Data Format: CSV
- Find data at: <https://www.bts.gov/latch/latch-data>

DATA DESCRIPTION:

- Pre-existing data collected by a third-party
- Modeling and simulation data: Input data and outputs for models such as CUBE, TRANSCad, or VISSIM; input data and output for simulation models created by PSR researchers.
- Contextual documentation: Included in project input folder

METADATA INFORMATION:

- Title: Local Area Transportation Characteristics for Households (LATCH Survey)
- Year: 2018
- Author: Bureau of Transportation Statistics, United States Department of Transportation

- Abstract: The purpose of the project was to develop estimates of average weekday household person trips, vehicle trips, person miles traveled, and vehicle miles traveled (per day), for all Census tracts in the United States.
- Methods: The Bureau of Transportation Statistics (BTS) developed a model that allows for Census tract estimation using the National Household Travel Survey (NHTS) data along with American Community Survey (ACS) data from the Census Bureau. The model divides the NHTS data into six geographic areas, classifies these areas as urban/suburban/rural, and then estimates average weekday household: person miles traveled, person trips, vehicle miles traveled, and vehicle trips for each geographic area. The BTS model then transfers the estimates to individual Census tracts using the household and demographic data from the ACS for each Census tract.
- Link to data source: <https://www.bts.gov/statistical-products/surveys/local-area-transportation-characteristics-households-latch-survey>

2.2. LODES Data (Lodes)

DATA STORAGE:

- Data Format: CSV and Shapefile
- Rational for Usage of Proprietary Data Format: The Shapefile format is open and popular for data transfer. An initial state format during map and shape digitization output, employed as a middle state format by many programs and publishers, and used for data transfer between GIS applications. Shapefiles can be created by exporting any data source to a shapefile, digitizing shapes directly, using programming software, or writing directly to the shapefile specifications by creating a program. Shapefiles are the simplest and most effective file format for geospatial information.
- Find data at: <https://lehd.ces.census.gov/data/>

DATA DESCRIPTION:

- Pre-existing data collected by a third-party
- Modeling and simulation data: Input data and outputs for models such as CUBE, TRANSCad, or VISSIM; input data and output for simulation models created by PSR researchers.
- Contextual documentation: Included in project input folder

METADATA INFORMATION:

- Title: LEHD Origin-Destination Employment Statistics
- Year: 2017
- Author: Longitudinal Employer-Household Dynamics, United States Census Bureau
- Abstract: A modeled data set utilizing administrative data from state partners to estimate the employment census block location for all workers living in each census block. For the simulation analysis, the data is aggregated to the Census Tract. Apart from the synthetic nature of the data, the limitation of this source is it will only provide data on work trips.
- Link to data source: <https://lehd.ces.census.gov/data/>

2.3. California Household Travel Survey (ca_travel_chts)

DATA STORAGE:

- Data Format: CSV
- Find data at: <https://www.nrel.gov/transportation/secure-transportation-data/tsdc-california-travel-survey.html>

DATA DESCRIPTION:

- Pre-existing data collected by a third-party
- Travel surveys: On-line or on-paper surveys administered to random samples of individuals or households as to daily travel patterns, attitudes and perceptions, and demographic characteristics. Datasets include individual and/or household records with names, addresses, and all other personal identifiers redacted.
- Contextual documentation: Included in project input folder

METADATA INFORMATION:

- Title: 2010–2012 California Household Travel Survey (CHTS)
- Author: Transportation Secure Data Center, National Renewable Energy Laboratory
- Year: 2017
- Abstract: The 2010–2012 California Household Travel Survey (CHTS) collected demographics and travel behavior characteristics for a multi-modal study of residents across the entire state. It provides the origin and destination tract for all trips with trip purpose information. At the time, it was the largest such regional or statewide survey ever conducted in the United States.
- Methods: Detailed travel behavior information was obtained from more than 42,500 households via multiple data-collection methods, including computer-assisted telephone interviewing, online and mail surveys, wearable (7,574 participants) and in-vehicle (2,910 vehicles) global positioning system (GPS) devices, and on-board diagnostic sensors that gathered data directly from a vehicle's engine. Details of personal travel behavior were gathered within the region of residence, inter-regionally within the state, and in adjoining states and Mexico. The survey sampling plan was designed to ensure an accurate representation of the entire population of the state. The CHTS included additional features—more detailed data on vehicle-acquisition decisions, parking choices, work schedules and flexibility, use of toll lanes/priced facilities, and walk and bicycle trips—to support advanced model development.
- Funding: Funded by the California Strategic Growth Council, the California Energy Commission (Energy Commission), and eight transportation planning agencies across the state.
- Link to data source: <https://www.nrel.gov/transportation/secure-transportation-data/tsdc-california-travel-survey.html>

NOTE:

"If you use TSDC data in a publication, please contact us and include a citation in your publication consistent with the following format:

"Transportation Secure Data Center." (2017). National Renewable Energy Laboratory. Accessed Jan. 15, 2017: www.nrel.gov/tsdc."

2.4. 2017 National Household Travel Survey - California (ca_travel_nhts_caaddon)

DATA STORAGE:

- Data Format: CSV and Shapefile
- Rational for Usage of Proprietary Data Format: The Shapefile format is open and popular for data transfer. An initial state format during map and shape digitization output, employed as a middle state format by many programs and publishers, and used for data transfer between GIS applications. Shapefiles can be created by exporting any data source to a shapefile, digitizing shapes directly, using programming software, or writing directly to the shapefile specifications by creating a program. Shapefiles are the simplest and most effective file format for geospatial information.
- Find data at: <https://www.nrel.gov/transportation/secure-transportation-data/tsdc-nhts-california.html>

DATA DESCRIPTION:

- Pre-existing data collected by a third-party
- Travel surveys: On-line or on-paper surveys administered to random samples of individuals or households as to daily travel patterns, attitudes and perceptions, and demographic characteristics. Datasets include individual and/or household records with names, addresses, and all other personal identifiers redacted.
- Contextual documentation: Included in project input folder

METADATA INFORMATION:

- Title: 2017 National Household Travel Survey – California Add-On
- Year: 2019
- Author: Transportation Secure Data Center, National Renewable Energy Laboratory
- Abstract: The California add-on survey supplements the 2017 National Household Travel Survey (NHTS) with additional household samples and detailed travel behavior for an assigned travel day. It provides travel diary information for a sample of households in California with origin and destination tracts available for all trip types.
- Methods: The Federal Highway Administration conducted the NHTS with assigned travel dates from April 19, 2016 through April 25, 2017. The NHTS collected data on the demographic and socioeconomic composition of households, as well as detailed information on travel behavior nationwide. State transportation departments and metropolitan planning agencies had the opportunity to purchase extra household samples as part of the NHTS add-on program. These additional samples, along with national samples collected in the add-on areas, are compiled for use in transportation planning, forecasting, and research. The California Department of Transportation participated in the NHTS add-on program and received a total of 26,095 household samples in California (total number of household samples nationwide is 129,112).
- Link to data source: <https://www.nrel.gov/transportation/secure-transportation-data/tsdc-nhts-california.html>

Note: If you use TSDC data in a publication, please contact us and include a citation in your publication consistent with the following format:

"Transportation Secure Data Center." (2019). National Renewable Energy Laboratory. Accessed Jan. 15, 2019: www.nrel.gov/tsdc.

2.5. Census Commuting Data

DATA STORAGE:

- Data Format: CSV
- Find Data at:
<https://data.census.gov/cedsci/table?q=ACSST1Y2019.B08301&g=0500000US06037.140000&tid=ACSDT5Y2017.B08301&hidePreview=true>

DATA DESCRIPTION:

- Pre-existing data collected by a third-party
- Demographic data: Data on individual and household characteristics as well as vehicle ownership and other travel-related characteristics acquired from state agencies and/or commercial vendors.
- Contextual documentation: Included in project input folder

METADATA INFORMATION:

- Title: Means of Transportation to Work (B08301)
- Year: 2017
- Author: US Census Bureau
- Abstract: The 5 year American Community Survey Table for commuting data aggregated at the Census Tract Level for tracts in Los Angeles County. Includes the estimate and margin of error for driving alone, carpooling, walking, and other modes of transportation to work. Used in this study to calculate the number of commuters in the projects Commuter Simulation.
- Link to Data Source:
<https://data.census.gov/cedsci/table?q=ACSDT1Y2019.B08301&g=0500000US06037.140000&tid=ACSDT5Y2017.B08301&hidePreview=true>

3. Freight Travel

3.1. HDT Model (SCAG)

DATA STORAGE:

- Data Format: Shapefile
- Rational for Usage of Proprietary Data Format: - Rational for Usage of Proprietary Data Format: The Shapefile format is open and popular for data transfer. An initial state format during map and shape digitization output, employed as a middle state format by many programs and publishers, and used for data transfer between GIS applications. Shapefiles can be created by exporting any data source to a shapefile, digitizing shapes directly, using programming software, or writing directly to the shapefile specifications by creating a program. Shapefiles are the simplest and most effective file format for geospatial information.
- Find data at: Data is proprietary and cannot be shared publicly. Data supplied by Southern California Association of Governments.

DATA DESCRIPTION:

- Pre-existing data collected by a third-party
- Modeling and simulation data: Input data and outputs for models such as CUBE, TRANSCad, or VISSIM; input data and output for simulation models created by PSR researchers.
- Contextual documentation: Included in project input folder under name 'scag_head.xlsx'

METADATA INFORMATION:

- Title: HEAVY-DUTY TRUCK MODEL
- Year: 2008
- Author: Southern California Association of Governments
- Abstract: Southern California Association of Governments (SCAG) developed the Heavy Duty Truck (HDT) Model to evaluate important policy choices and investment decisions. The HDT model is a primary analysis tool to support the goods movement policy decisions made by SCAG and regional stakeholders.
- Link to data source: <https://scag.ca.gov/heavy-duty-truck-model>

4. Geographic Data Sources

4.1. ACS and OSMnx (tracts_indicators_grades_eras_index.csv)

DATA STORAGE:

- Data Format: CSV

DATA DESCRIPTION:

- Pre-existing data collected by a third-party
- Data Description - Demographic data: Data on individual and household characteristics as well as vehicle ownership and other travel-related characteristics acquired from state agencies and/or commercial vendors.
- Contextual documentation: Included in project input folder

METADATA INFORMATION:

- Title: Off the Grid...and Back Again?
- Year: 2020
- Author: Geoff Boeing, boeing@usc.edu, University of Southern California
- Abstract: In this morphological study I identify and measure recent nationwide trends in American street network design. Historically, orthogonal street grids provided the interconnectivity and density that researchers identify as important factors for reducing vehicular travel and emissions and increasing road safety and physical activity. During the 20th century, griddedness declined in planning practice along with declines in urban form compactness, density, and connectivity as urbanization sprawled around automobile dependence. But less is known about comprehensive empirical trends across U.S. neighborhoods, especially in recent years. Here I use public and open data to examine tract-level street networks across the entire United States. I develop theoretical and measurement frameworks for a quality of street networks defined here as griddedness. I measure how griddedness, orientation order, straightness, 4-way intersections, and intersection density declined from 1940 through the 1990s, while dead-ends and block lengths increased. However, since 2000, these trends have rebounded, shifting back toward historical design patterns. Despite this rebound, when controlling for topography and built environment factors, all decades after 1939 are associated with lower griddedness than pre-1940 decades. Higher griddedness is associated with less car ownership—which itself has a well-established relationship with vehicle kilometers traveled and greenhouse gas emissions—while controlling for density, home and household size, income, jobs proximity, street network grain, and local topography.
- Link to data source: <https://www.tandfonline.com/doi/full/10.1080/01944363.2020.1819382>

4.2. OpenStreetMap

DATA STORAGE:

- Web

DATA DESCRIPTION:

- Pre-existing data collected by a third-party
- Data Description: OpenStreetMap (OSM) is a collaborative project to create a free editable map of the world. The data pulled from OSM for this project was a snapshot of the street network in Los Angeles County on March 4, 2021. OpenStreetMap is an open source dataset, and you can find the copyright page at the link below:
Link: <https://www.openstreetmap.org>

METADATA INFORMATION:

- Title: OpenStreetMap
- Year: 2021
- Author: © OpenStreetMap contributors

4.3. Census Block ShapeFiles (blocks_shapefile10)

DATA STORAGE:

- Data Format: Shapefile
- Rational for Usage of Proprietary Data Format: - Rational for Usage of Proprietary Data Format: The Shapefile format is open and popular for data transfer. An initial state format during map and shape digitization output, employed as a middle state format by many programs and publishers, and used for data transfer between GIS applications. Shapefiles can be created by exporting any data source to a shapefile, digitizing shapes directly, using programming software, or writing directly to the shapefile specifications by creating a program. Shapefiles are the simplest and most effective file format for geospatial information.
- Find data at: <https://www.census.gov/geographies/mapping-files/time-series/geo/tiger-line-file.2010.html>

DATA DESCRIPTION:

- Pre-existing data collected by a third-party
- Vector Spatial Data: Spatial data comprise the relative geographic information about the earth and its features. A pair of latitude and longitude coordinates defines a specific location on earth. Spatial data are of two types according to the storing technique, namely, raster data and vector data. Raster data are composed of grid cells identified by row and column. The whole geographic area is divided into groups of individual cells, which represent an image. Vector data are composed of points, polylines, and polygons.
- Contextual documentation: Included in project input folder

METADATA INFORMATION:

- Title: TIGER/Line Shapefiles
- Year: 2010
- Author: Spatial Data Collection and Products Branch, United States Census Bureau
- Abstract: The TIGER/Line Shapefiles are extracts of selected geographic and cartographic information from the Census Bureau's Master Address File (MAF)/Topologically Integrated Geographic Encoding and Referencing (TIGER) Database (MTDB). The shapefiles include information for the 50 states, the District of Columbia, Puerto Rico, and the Island Areas (American Samoa, the Commonwealth of the Northern Mariana Islands, Guam, and the United States Virgin Islands). The shapefiles include polygon boundaries of geographic areas and features, linear features including roads and hydrography, and point features. These shapefiles do not contain any sensitive data or confidential data protected by Title 13 of the U.S.C.
- Link to data source: <https://www.census.gov/geographies/mapping-files/time-series/geo/tiger-line-file.2010.html>

4.4. Tracts Shapefile (tracts_shapefile)

DATA STORAGE:

- Data Format: Shapefile
- Rational for Usage of Proprietary Data Format: - Rational for Usage of Proprietary Data Format: The Shapefile format is open and popular for data transfer. An initial state format during map and shape digitization output, employed as a middle state format by many programs and publishers, and used for data transfer between GIS applications. Shapefiles can be created by exporting any data source to a shapefile, digitizing shapes directly, using programming software, or writing directly to the shapefile

specifications by creating a program. Shapefiles are the simplest and most effective file format for geospatial information.

- Find data at: <https://www.census.gov/geographies/mapping-files/time-series/geo/tiger-line-file.2019.html>

DATA DESCRIPTION:

- Pre-existing data collected by a third-party

- Vector Spatial Data: Spatial data comprise the relative geographic information about the earth and its features. A pair of latitude and longitude coordinates defines a specific location on earth. Spatial data are of two types according to the storing technique, namely, raster data and vector data. Raster data are composed of grid cells identified by row and column. The whole geographic area is divided into groups of individual cells, which represent an image. Vector data are composed of points, polylines, and polygons.

- Contextual documentation: Included in project input folder

METADATA INFORMATION:

- Title: TIGER/Line Shapefiles

- Year: 2017

- Author: Spatial Data Collection and Products Branch, United States Census Bureau

- Abstract: The TIGER/Line Shapefiles are extracts of selected geographic and cartographic information from the Census Bureau's Master Address File (MAF)/Topologically Integrated Geographic Encoding and Referencing (TIGER) Database (MTDB). The shapefiles include information for the 50 states, the District of Columbia, Puerto Rico, and the Island Areas (American Samoa, the Commonwealth of the Northern Mariana Islands, Guam, and the United States Virgin Islands). The shapefiles include polygon boundaries of geographic areas and features, linear features including roads and hydrography, and point features. These shapefiles do not contain any sensitive data or confidential data protected by Title 13 of the U.S.C.

- Link to data source: <https://www.census.gov/geographies/mapping-files/time-series/geo/tiger-line-file.2019.html>

4.5. Freeway ShapeFiles (Freeway)

DATA STORAGE:

- Data Format: CSV and Shapefile

- Rational for Usage of Proprietary Data Format: - Rational for Usage of Proprietary Data Format: The Shapefile format is open and popular for data transfer. An initial state format during map and shape digitization output, employed as a middle state format by many programs and publishers, and used for data transfer between GIS applications. Shapefiles can be created by exporting any data source to a shapefile, digitizing shapes directly, using programming software, or writing directly to the shapefile specifications by creating a program. Shapefiles are the simplest and most effective file format for geospatial information.

- Find data at: <https://www.census.gov/cgi-bin/geo/shapefiles/index.php>

DATA DESCRIPTION:

- Pre-existing data collected by a third-party

- Vector Spatial Data: Spatial data comprise the relative geographic information about the earth and its features. A pair of latitude and longitude coordinates defines a specific location on earth. Spatial data are of two types according to the storing technique, namely, raster data and vector data. Raster data are composed of grid cells identified by row and column. The whole geographic area is divided into groups

of individual cells, which represent an image. Vector data are composed of points, polylines, and polygons.

Contextual documentation: Included in project input folder

METADATA INFORMATION:

- Title: TIGER/Line Shapefiles

- Year: 2013

- Author: Spatial Data Collection and Products Branch, United States Census Bureau

- Abstract: The TIGER/Line Shapefiles are extracts of selected geographic and cartographic information from the Census Bureau's Master Address File (MAF)/Topologically Integrated Geographic Encoding and Referencing (TIGER) Database (MTDB). The shapefiles include information for the 50 states, the District of Columbia, Puerto Rico, and the Island Areas (American Samoa, the Commonwealth of the Northern Mariana Islands, Guam, and the United States Virgin Islands). The shapefiles include polygon boundaries of geographic areas and features, linear features including roads and hydrography, and point features. These shapefiles do not contain any sensitive data or confidential data protected by Title 13 of the U.S.C.

- Link to data source: <https://www.census.gov/cgi-bin/geo/shapefiles/index.php>

- Link to data source: <https://www.census.gov/cgi-bin/geo/shapefiles/index.php>

4.6. SCAG Open Spaces (openspaces)

DATA STORAGE:

- Data Format: Shapefile

- Rational for Usage of Proprietary Data Format: - Rational for Usage of Proprietary Data Format: The Shapefile format is open and popular for data transfer. An initial state format during map and shape digitization output, employed as a middle state format by many programs and publishers, and used for data transfer between GIS applications. Shapefiles can be created by exporting any data source to a shapefile, digitizing shapes directly, using programming software, or writing directly to the shapefile specifications by creating a program. Shapefiles are the simplest and most effective file format for geospatial information.

- Find data at: [https://gisdata-](https://gisdata-scag.opendata.arcgis.com/datasets/c2c9a1a6e5064302aff5631812ecbeb7_0/data)

[scag.opendata.arcgis.com/datasets/c2c9a1a6e5064302aff5631812ecbeb7_0/data](https://gisdata-scag.opendata.arcgis.com/datasets/c2c9a1a6e5064302aff5631812ecbeb7_0/data)

DATA DESCRIPTION:

- Pre-existing data collected by a third-party

- Vector Spatial Data: Spatial data comprise the relative geographic information about the earth and its features. A pair of latitude and longitude coordinates defines a specific location on earth. Spatial data are of two types according to the storing technique, namely, raster data and vector data. Raster data are composed of grid cells identified by row and column. The whole geographic area is divided into groups of individual cells, which represent an image. Vector data are composed of points, polylines, and polygons.

- Contextual documentation: SCAG did not create these documents

METADATA INFORMATION:

- Title: Protected Open Space Areas: SCAG Regions

- Year: 2021

- Author: Southern California Association of Governments

- Abstract: This dataset includes the boundaries of California Protected Areas, lands that are protected for open space purposes. Data includes all such areas in California, from small urban parks to large national parks and forests, mostly aligned to assessor parcel boundaries. Data is collected by Holdings

(parcels) which are aggregated to Units (commonly named areas within a county) and Super Units (commonly named areas generally).

- Link to data source: [https://gisdata-](https://gisdata-scag.opendata.arcgis.com/datasets/c2c9a1a6e5064302aff5631812ecbeb7_0/data)

[scag.opendata.arcgis.com/datasets/c2c9a1a6e5064302aff5631812ecbeb7_0/data](https://gisdata-scag.opendata.arcgis.com/datasets/c2c9a1a6e5064302aff5631812ecbeb7_0/data)

Data Access and Sharing

These data are available as described above from their original creators.

Reuse and Redistribution

Restrictions on reuse and redistribution are defined by each dataset's original creator as defined above.