ECONOMICS 620

Allegheny College
Meadville, Pennsylvania 16335

SUSTAINABILITY COST OF AUTOMOBILE DEPENDENCY IN 100 AMERICAN CITIES

Ryan Sheehan

April 6, 2025

SUSTAINABILITY COST OF AUTOMOBILE DEPENDENCY IN 100 AMERICAN CITIES

by

Ryan Sheehan

Submitted to The Department of Business and Economics

Project Advisor: Francisco Navarro-Sanchez

Second Reader: Chris Finaret

Date: April 6, 2025

Acknowledgements

Thank you to my project advisor Professor Navarro-Sanchez and to Professor Finaret for their feedback and guidance during my composition of this paper. Thank you to Professor Nonnenmacher, my academic advisor, and to the rest of the business and economics faculty who have made my time at Allegheny fulfilling.

Table of Contents

List of Tables	v
List of Figures	vi
Summary of AI Use by Author	vii
Abstract	viii
Chapter 1: Introduction	1
Chapter 2: Literature Review	5
Chapter 3: Methodology	11
Chapter 4: Results and Discussion	19
Chapter 5: Conclusion	37
References	40

List of Tables

Table 3.1: Variables and sources

Table 3.2: Variables and the geographic areas from which they were sampled

Table 4.1: Results for traffic fatalities (density dummies, all cities)

Table 4.2: Results for traffic fatalities (average density, all cities)

Table 4.3: Results for traffic fatalities (density dummies, filtered cities)

Table 4.4: Results for traffic fatalities (average density, filtered cities)

Table 4.5: Results for median AQI (density dummies)

Table 4.6: Results for median AQI (average density)

Table 4.7: Results for percentage of days with AQI<50 (density dummies)

Table 4.8: Results for percentage of days with AQI<50 (average density)

Table 4.9: Results for vehicles per household (density dummies)

Table 4.10: Results for vehicles per household (average density)

Table 4.11: Results for median AQI with vehicles per household as an IV (density dummies)

Table 4.12: Results for median AQI with vehicles per household as an IV (average density)

Table 4.13: Results for percentage of days with AQI<50 with vehicles per household as an IV (density dummies)

Table 4.14: Results for percentage of days with AQI<50 with vehicles per household as an IV (density dummies)

List of Figures

Figure 4.1: Distribution of Traffic Fatalities per 100,000 People

Figure 4.2: Distribution of P>12,500

Figure 4.3: Distribution of P<1,500

Figure 4.4: Distribution of Vehicle Miles Travelled Per Capita

Figure 4.5: Correlation Between VMT and Average Density

Figure 4.6: Correlation Between VMT and Proportion of Commuters who Don’t Drive

Figure 4.7: Distribution of Vehicles per Household

Figure 4.8: Distribution of Population Density

Summary of AI Use by Author

No artificial intelligence tools were used for this research.

Abstract

I used cross-sectional data from the 100 most populous cities in the United States to determine the impact of automobile dependency on social, environmental, and economic sustainability outcomes. This research departed from previous literature by using data from cities rather than states and by incorporating urban sprawl as a measure of automobile dependency. Sustainability outcomes were quantified using traffic fatalities per 100,000 people as a measure of social costs, air quality as a measure of environmental costs, and vehicle ownership as a measure of economic costs. Measures of automobile dependency included vehicle miles traveled (VMT) per capita, three measures of population density, and two measures of commute mode. My findings indicate that VMT per capita is associated with increased traffic fatalities and increased vehicle ownership. The proportion of commuters who do not drive is associated with lower traffic fatalities, better air quality, and fewer vehicles per household. Contrary to my prediction, the proportion of drivers who carpool is associated with increased traffic fatalities. Measures of population density yielded conflicting results.

Chapter 1: Introduction

Automobile dependency is becoming a mainstream talking point as the world begins to cut its greenhouse gas emissions and focus on sustainability. Automobiles are the least efficient mode of transportation in terms of the amount of space taken up by each commuter, and the area devoted to them is largely responsible for the degradation and decentralization of urban environments around the world. The United States is a particularly notable offender. Americans' vehicle use has more than doubled since 1980 to over three trillion vehicle miles traveled (VMT) per year (Office of Highway Policy Information, 2022). While the traditional measure of a transportation network primarily assesses mobility, greater attention is being paid to measures of sustainability. Recent research has emphasized negative effects of automobile dependency that go beyond conventional measures of congestion, throughput, and efficiency. My research will focus on the effect of automobile dependency on social, environmental, and economic sustainability outcomes in the United States' largest metropolitan statistical areas (MSAs).

In order to quantify these effects, I used cross-sectional data from the 100 largest MSAs in the United States. The independent variable, automobile dependency, will be measured using data on driving habits and vehicle ownership, as well as factors related to urban sprawl. Dependent variables will include sustainability measures such as air quality, traffic fatalities, and transportation spending by households.

The framework for this research was pioneered by Zheng et al. (2011), who proposed sustainability measures based on the Canadian Center for Sustainable Transport's three domains of sustainability: environmental sustainability, social sustainability, and economic sustainability. These domains were then measured at the state level along with VMT. Studies using this framework have found that lowering a state's VMT can significantly mitigate the effects of some

of the most serious externalities of automobile dependency such as pollution and traffic fatalities (Ahangari et al., 2017, Garceau et al., 2013, Zheng et al., 2011).

These studies also reveal that the effects of automobile use are most extreme in rural areas. For instance, Wyoming is the least populous and one of the most rural states and was found by Garceau et al. (2013) to have the highest VMT per capita, government transportation spending per capita, carbon dioxide emissions from transportation per capita, and highest rates of deaths in vehicle crashes per capita, which I will be referring to for the rest of this paper as traffic fatalities. This is similar to the findings of Ahangari et al. (2017) that the proportion of a state living in rural areas is significantly associated with greater traffic fatalities, and the density of a state's densest area is significantly associated with fewer. This suggests that the effects of vehicle use fall rather neatly along the urban/rural divide. Because of this, variables that are meant to measure automobile dependency and its effects could simply be a function of the proportion of a state that lives in urban versus rural areas; the higher the proportion of the population that lives in a rural area, the worse that state's sustainability outcomes will be. Employing data from American MSAs should mitigate this effect and provide a better understanding of the effects of automobile dependency on urban areas, the difference between an urban area that is automobile dependent and one that is not, and what can be done to combat automobile dependency in urban areas.

Another trend associated with automobile dependency is urban sprawl. Sprawl is the result of horizontal rather than vertical expansion of urban environments. It leads to lower residential density as well as loss of natural habitats and farmland. It also perpetuates a cycle wherein homes are built farther and farther from the destinations their residents need to access, requiring residents to drive and increasing dependency on automobiles. Increased automobile use

makes areas even farther from destinations more accessible, thus exacerbating urban sprawl in what Silva and Altieri (2022) call a “feedback loop”.

Ewing et al. (2003) attempt to quantify urban sprawl by measuring variables associated with residential density, a measure of how developed the city is, on average; land use mix, the density of non-residential destinations; degree of centering, the amount density increases near the city center; and street accessibility, the density of the street network.

Because urban sprawl and automobile use are so intertwined, it would stand to reason that the negative effects of urban sprawl are also effects of automobile dependency, and vice versa. Urban sprawl itself could be classified as a proxy for automobile dependency. Because of this, an analysis of a transportation system based on sustainability should also take into account the level of urban sprawl in MSAs and its effects. For my research, I combined the framework developed by Zheng et al. (2011) with the measures of urban sprawl developed by Ewing et al. (2003) to evaluate the effects of urban sprawl on the three domains of sustainability.

In order to quantify sustainability outcomes, I used the measures employed by Garceau et al. (2013), but with city level rather than state level data. I employed traffic fatalities per 100,000 people as a measure of social costs. I used a city’s median Air Quality Index (AQI) and the percentage of days where a city’s AQI was below 50 as measures of environmental costs. For economic factors, I used each city’s average number of vehicles per household.

My independent variables are measures of automobile dependency. I used VMT per capita as well as commuting data to test the effect of the proportion of the population traveling to work in a privately owned vehicle versus any other method such as walking or public transportation. These variables will quantify driving habits, which is a more conventional measure of automobile dependency.

I also employed urban sprawl as a measure of automobile dependency. Ewing et al. (2003) breaks urban sprawl into four categories: residential density, land use mix, degree of centeredness, and street accessibility. Each of these categories receives an aggregate score which is calculated using several different variables. Rather than calculating composite scores, I simply used some of the variables used in the calculation. Density is the most obvious measure of urban sprawl, so that is what I focused on for this research. I used the average population per square mile, proportion of population living at a density of 1,500 people per square mile or less, and the proportion of the population living at a density of 12,500 people per square mile or more as measures of residential density. I excluded the other three categories, firstly because data was difficult to find, and secondly because neither of the latter two was found to be statistically significant (Ewing et al., 2003).

My results reflected the findings of similar past research. Increased VMT was associated with greater traffic fatality rates and more vehicles per household. Increased density was associated with lower traffic fatalities. Greater travel by any means other than a motor vehicle was associated with lower traffic fatalities, better air quality, and fewer vehicles per household. Unlike other literature, however, and contrary to my expectations, carpooling was associated with greater traffic fatality rates and higher density was associated with more vehicles per household and lower air quality.

Chapter 2: Literature Review

Automobiles are an integral part of any transportation network. However, transportation infrastructure and policy across the globe, and especially in the United States, relies heavily on automobile use as the only mode of transportation. This dependency decreases the efficiency and increases the cost of maintaining road networks, ultimately placing too high a burden on the automobile as a monomodal solution to transport. Roads in the United States no longer provide the high marginal benefits that they once did. Instead, drivers suffer from inefficient allocation of transportation resources (Shirley & Winston, 2004), longer commutes to access the same destinations (Kooshian & Winkelman, 2011), and increased congestion (Noland & Lem, 2002).

Shirley and Winston (2004) use firm inventories as a measure of efficiency for highway infrastructure. Increased mobility allows firms to store fewer goods and decreases inventory and transportation costs. Using this framework, they measure the efficiency of transportation infrastructure spending over time. They find that highway spending in the 1970s had a 17.6% return as measured by inventory and transportation savings; each dollar spent on highway infrastructure saved businesses 17.6 cents. This figure fell to 4.9% in the 1980s and 1.0% in the 1990s. This could be attributed to a shift in focus from highway construction to highway maintenance, or to inefficient allocation of funds. Additionally, Kooshian and Winkelman (2011) discuss the relationship between vehicle miles traveled (VMT) and prosperity. They contend that the low efficiency of automobile use is a result of the layout of cities and road networks and a lack of alternative modes of transportation, and that this leads to drivers needing to take longer trips to access the same destinations. They call extra VMT accrued for these reasons “empty miles”.

Freeway expansion also suffers from decreased efficiency due to the law of induced demand. If the price of driving is measured as time in traffic and supply as the capacity of a transportation network, then increasing supply leads to a decreased price and higher demand. The result is that expanded highway infrastructure actually leads to increased congestion and lower mobility (Noland & Lem, 2002).

While congestion is one of the most obvious externalities of automobile use, the majority of gripes with automobile dependency do not stem from concerns about efficiency and spending but from concerns about the environment and public health. Automobile use comes with many negative externalities that are not addressed by the relatively low cost of driving or by US transportation policy. Edlin and Karaca-Mandic (2006) write about the monetary cost of automobile crashes. They suggest that, because drivers choose automobile travel over another mode of transportation, both parties involved in an automobile crash are responsible for the full cost of the crash. However, the financial cost is only paid by one party. Thus, only half of the social cost of the crash is accounted for. Other economic issues stem from automobile inefficiency as well. Kang (2021) finds that automobile dependency increases rates of housing insecurity. Families that do not own a car are 45% more likely to experience housing instability. This figure increases to 85% in car dependent metropolitan statistical areas (MSAs).

The externalities of automobile dependency go beyond monetary factors. Frederick, Riggs, and Gilderbloom (2018) address public health effects of car dependency, finding that metropolitan regions with lower rates of commuting by single occupancy vehicles have lower rates of obesity and higher levels of physical activity. Similarly, Frank, Andresen, and Schmid (2004) find a significant relationship between urban form and weight. They measure urban form using spatial analysis models which incorporate land use and intersection density. Their research

finds that the likelihood of obesity increases by 6.0% for each hour per day spent in a car, and decreases by 4.8% for each kilometer walked per day.

One of the most detrimental factors to the quality of urban form is sprawl. Urban sprawl decreases the population density of cities and increases the space they take up, leading to the loss of farmland and natural habitats. Silva and Altieri (2022) explore the relationship between automobile use and urban sprawl. Their research finds that local accessibility, as measured by destinations within a fifteen-minute walking distance, is being sacrificed for regional accessibility, measured by destinations within a fifteen-minute driving distance. This means people are living farther away from where they need to go, which perpetuates car use, which then encourages sprawl in a "feedback loop". In this way, automobile use and urban sprawl are results of the same issues and trends.

Urban sprawl is a difficult phenomenon to quantify. Metropolitan areas do not have well defined borders because there is rarely a clear cutoff between urban and rural areas. Measures often overestimate or underestimate density by including too little or too much land (Wolman et al., 2005). Nonetheless, it is an important characteristic to define. Ewing et al. (2003) develops a detailed system for defining urban areas and measuring sprawl by breaking it into four separate measures: residential density, land use mix, degree of centering, and street accessibility. These measures are then tested against various transportation outcomes. Most significantly, findings indicate that higher density leads to lower rates of vehicle ownership, vehicle miles traveled (VMT), and traffic fatalities, as well as higher rates of public transportation use. Centeredness was found to have a similar effect. Stone (2008) uses the measures of urban sprawl developed by Ewing et al. (2003) to create a composite value for urban sprawl and test its effect on air quality. Air quality is measured using ozone exceedances as a proxy because they are more localized

than other greenhouse gases. The research shows that MSAs that exhibit higher levels of urban sprawl have statistically significantly more ozone exceedances, which suggests that urban sprawl leads to lower air quality.

Policies have been implemented across the nation to combat urban sprawl and its adverse effects. Zhu et al. (2022) discuss the efficacy of these policies. They find that investment in transportation infrastructure and regulation of land use leads to a statistically significant increase in the number of people carpooling or using public transportation. They also find that younger generations are more likely to be responsive to these changes. Similarly, Cervero and Murakami (2010) assess the effect of land use patterns on VMT in 370 American MSAs. They measure land use primarily based on population density and accessibility of destinations. Findings indicate that increases in either of these measures lead to a decrease in VMT, and that destination accessibility has a stronger impact on VMT. The effect of density is offset slightly by the fact that increased density is associated with higher road density and more abundant destinations, both of which can lead to increases in VMT.

One of the most serious and easily quantifiable externalities of automobile dependency is automobile crashes. Traffic fatalities are explored by Ahangari et al. (2017) by analyzing their response to different factors related to driving. Among the factors included are the price of fuel, vehicle ownership rates, proportion of people walking to work, density, and healthcare quality. Most notably, the authors find a statistically significant, negative relationship between traffic fatalities and the percentage of the population that walks to work, and between traffic fatalities and the density of a state's densest area, as well as a positive relationship between traffic fatalities and the percent of a state's population living in low-density, rural areas.

One reason increased driving can lead to more dangerous road networks is an "arms race" among drivers, particularly in the United States. In the case of an automobile crash, drivers of larger vehicles are less likely to be seriously injured or killed than drivers in smaller vehicles. However, drivers of smaller vehicles are more likely to be seriously injured or killed when involved in a crash with a larger vehicle. Drivers prefer to buy larger vehicles despite the negative impact on society because it is safer for the driver (White, 2004). The result is that American roads are more deadly than those of other developed countries.

Zheng et al. (2011) incorporate externalities of automobile use into their measure of a transportation network. Where prior research focuses primarily on measures such as efficiency and throughput, the authors of this study suggest that sustainability, which they divide into environmental, social, and economic domains, is an equally important part of measuring a transportation network. The researchers focus on economic sustainability, which they measure using variables including the proportion of household income spent on transportation, state gross domestic product (GDP) compared to state VMT, state GDP growth compared with VMT growth, how much of a state's transportation spending comes from federal funding, and how much of a state's GDP is spent on fuel. The authors find that states' outcomes depend on where they fall on the urban/rural divide. States with higher levels of urbanity perform better than rural states in terms of the economic sustainability of their transportation networks.

The framework constructed by Zheng et al. (2011) is used by Garceau et al. (2013) to examine all three domains of sustainability. Carbon dioxide emissions per capita are used to measure environmental sustainability, traffic fatalities are used to measure social sustainability, and transportation spending both by the government and by households are used to measure economic sustainability. The authors find that higher VMT is significantly associated with higher

carbon dioxide emissions, higher rates of traffic fatalities, and higher levels of spending both by households and states. Higher transportation spending by states is likely due to the high cost of maintaining road networks and inefficient allocation of transportation funding. Also of note is the fact that traffic fatalities increase with VMT at a nonlinear rate, meaning increased VMT leads to increased likelihood of a fatality per mile.

Chapter 3: Methodology

I am using cross-sectional data from 2021 for this study. Prior research in the literature uses statewide data to assess the effects of automobile dependency in each state (Zheng et al., 2011, Garceau et al., 2013, Ahangari et al., 2017). As I have discussed, the downside of this approach is that outcomes are determined simply by the level of urbanization of each state; Montana and Wyoming had higher levels of measured automobile dependency and higher sustainability costs than Massachusetts and Rhode Island. One benefit of the statewide approach is that each state has clearly defined borders and data from within these borders is readily available. Cities are more difficult to study because they are not so clearly defined. Data from within the borders of an actual municipality is not enough; cities are not self-contained, but have a significant impact on their surroundings. It would be misleading to look at cities like Boston and Cambridge, Massachusetts, for example, as two entirely separate entities. Instead, we need some way of deciding which areas surrounding a central municipality count as part of that city.

Two definitions of a city are important for this research. A Core-based statistical area (CBSA) is made up of counties that are economically integrated with a central city with a population of at least 10,000. A county is determined to be economically integrated with a city if a certain proportion of the county's population commutes to the city for work. A metropolitan statistical area (MSA) is a type of CBSA whose central city has a population of at least 50,000. Because I am focusing only on major cities, an MSA and a CBSA will be synonymous for the purposes of this research.

Consider a county where the majority of the population lives in one town that is economically integrated with a nearby city but is otherwise sparsely populated. This county would be considered part of that city's MSA despite being mostly rural. A federal-aid urbanized

area (which I will simply be calling an urbanized area or an actual urbanized area) includes contiguous census tracts with a combined total of at least 2,000 housing units or a population of 5,000, all of which must meet a certain housing density threshold. In 2020, this threshold was 200 units per square mile (United States Census Bureau, 2022). Census tracts are much smaller than counties, so this method delineates a more precise border and excludes rural areas.

In keeping with the research of Garceau et. al (2013), I collected data for independent variables related to automobile dependency, including vehicle miles travelled (VMT) per capita and commute mode data. VMT for the 105 most populated urbanized areas in the United States (not including Puerto Rico) as well as their populations came from the Federal Highway Administration Highway Statistics Series Table HM-71 and was used to calculate annual VMT per capita for each city (Office of Highway Policy Information, 2023).

Data that was organized by individual census tracts needed to be filtered to only include tracts that were part of a given city's urbanized area. To do this, I filtered the data I collected for each census tract to include only tracts from the counties that made up that city's metropolitan area. I then filtered out the tracts that did not meet a density of 200 housing units per square mile, which was the 2020 threshold for inclusion in a given city's urbanized area. This method should leave behind every census tract that is included in the urbanized area, as well as some tracts that are not part of the urbanized area because they are not adjacent. I will be referring to this area as the estimated urbanized area. I have no way to filter out non-contiguous tracts. As a result, the population of the estimated urbanized area using this method slightly exceeded the actual population of the urbanized area in most cases. I calculated the population of each MSA using the same method, but without filtering based on population density.

As an example of how this method approximates the urbanized area, consider the city of Los Angeles. The Los Angeles urbanized area is the second most populous in the United States, with a census population of 12,150,996. The Los Angeles MSA contains two counties, Los Angeles County and Orange County, California. Together, these counties contain 3,112 census tracts with a total population of 13,200,998. After filtering out census tracts with a density lower than 200 housing units per square mile, the Los Angeles estimated urbanized area contains 3,010 census tracts with a total population of 12,980,069.

Using this methodology, I calculated commuting data and population density data, both independent variables, as well as vehicle ownership, which will be my proxy for economic sustainability. Variables representing commuting data include the proportion of commuters who commute in a motor vehicle and the proportion who use another mode of transportation such as public transportation, walking, or cycling. I also calculated the proportion of motor vehicle commuters who carpool versus the proportion who drive alone. All of these values were calculated using the total number of commuters using a given mode in each census tract included in the estimated urbanized area. This data came from the Bureau of Transportation Statistics' Intersections Data and Equity Hub (Bureau of Transportation Statistics).

I calculated the population density of each census tract using its 2021 census population, obtained from the United States Census Bureau (2020), and its geographic area, obtained from the Census Bureau's TIGER Shapefiles (United States Census Bureau, 2021). I divided the combined area of every census tract in each estimated urbanized area by the total population of each estimated urbanized area to calculate the average population density per square mile for each urbanized area. I also calculated the proportion of the population living in a census tract with a density greater than 12,500 people per square mile for each urbanized area, and the

proportion of the population living in a census tract with a density lower than 1,500 people per square mile.

Vehicle ownership data as well as the total number of households in each census tract came from the same source as commute mode data. I used the number of households in each census tract owning one, two, three, or four or more vehicles to calculate the total number of vehicles in each census tract. Households in the 'four or more' category were counted as owning exactly four vehicles, so this likely underestimated the true count slightly. I used the sum of the number of vehicles in every census tract comprising an estimated urbanized area, as well as the total number of households to determine the average number of vehicles per household for each estimated urbanized area. Similar research in the literature multiplied this value by the American Automobile Association's average annual cost of owning a motor vehicle in order to approximate household transportation spending (Zheng et. al, 2011, Garceau et. al, 2013). However, because this cost is constant, I will simply use the average number of vehicles per household as my economic variable.

Other dependent variables include proxies for the three domains of sustainability discussed by Zheng et al (2011). As a measure of social sustainability, I collected data on traffic fatalities from the National Highway Traffic Safety Administration's Fatality Analysis Reporting System, which listed every fatality from 2021 and where it occurred (National Highway traffic Safety Administration, 2024). Data on each traffic fatality contained which county it took place in but not which census tract. As such, I simply calculated the total number of traffic fatalities that occurred in counties comprising a given MSA and used the total population of that MSA to determine the number of traffic fatalities per 100,000 people.

As a measure of environmental sustainability, I collected Air Quality Index (AQI) data for each city from the United States Environmental Protection Agency (2021). AQI data was not available for five of the 105 most populous urbanized areas, so these were left out of the sample, bringing the sample size for this research to an even 100. AQI data was only available for MSAs as the air quality of individual census tracts is not monitored. However, I would expect air quality to remain relatively consistent despite slight changes in the geographic area from which it is collected. I used this data to calculate two separate measures of air quality; the median AQI for every day where it was monitored, and the proportion of days where AQI was monitored and below 50. An AQI of 50 or lower is considered 'good', so this variable is the proportion of days where air quality in a given city was good.

Table 3.1: ***Variables and Sources***

Variable	Source
Vehicle Miles Travelled (2021)	Office of Highway Policy Information. (2023). *Highway Statistics 2021 Table HM-71.*
Population by Urbanized Area (2021)	Office of Highway Policy Information. (2023). *Highway Statistics 2021 Table HM-71.*
Population by Census Tract (2020)	United States Census Bureau. (2020). *Total Population.* .
Housing Units by Census Tract (2020)	United States Census Bureau. (2020). *Housing*
Commute Mode (2021)	Bureau of Transportation Statistics.*Vehicle Ownership and Commuting Most Recent (2021) Selected Variables.*
Number of Vehicles Per Household (2021)	Bureau of Transportation Statistics.*Vehicle Ownership and Commuting Most Recent (2021) Selected Variables.*
Geographic Area by Census Tract (2021)	United States Census Bureau. (2021). *TIGER Shapefiles.*
Traffic Fatalities (2021)	National Highway Traffic Safety Administration. (2024). *Fatality Analysis Reporting System.*
Air Quality Index (2021)	United States Environmental Protection Agency. (2021). *Air Quality Statistics Report.*

Table 3.2: ***Variables and the geographic areas from which they were sampled***

Independent Variables	Area
Vehicle Miles Travelled Per Capita	Actual Urbanized Area
Commute Mode	Estimated Urbanized Area
Population Density	Estimated Urbanized Area
Dependent Variables	
Traffic Fatalities	Metropolitan Statistical Area
Air Quality Index	Metropolitan Statistical Area
Number of Vehicles Per Household	Estimated Urbanized Area

The distinct differences between the areas from which my variables are derived will likely have a slight impact on their values. I expect the extra space included within the estimated urbanized area but not the actual urbanized area to display characteristics that would be associated with more rural areas. As such, I expect commute mode data to lean slightly more towards motor vehicle use, I expect the number of vehicles per household to be slightly higher, and I expect population density to be slightly lower in estimated urbanized areas than in actual urbanized areas. In the case of population density, this would manifest as a lower average score for the city, as well as a lower proportion living at a density higher than 12,500 people per square mile and a higher proportion living at a density lower than 1,500 people per square mile. All of this would be in keeping with the findings of Zheng et. al (2011) and Garceau et. al (2013), who observed higher levels of car dependency in more rural samples. However, these differences should be relatively small, and underestimates and overestimates should be fairly consistent across cities. As such, I do not expect this to interfere with the results of this research.

Traffic fatalities and AQI are based on MSA rather than urbanized area or estimated urbanized area. I expect that AQI will remain practically identical no matter which method is

used due to the fact that air quality is not something that is likely to vary widely across small geographic areas. Traffic fatalities had to be collected by MSA because location data was not precise enough to rule out cases that may have occurred in more sparsely populated areas, meaning the number of fatal crashes in urbanized areas was probably overestimated. In order to account for this fact, I calculated the number of fatalities per 100,000 residents using the population of each city's MSA, which is simply the sum of the population of each county that comprises the MSA. This is the same sample that traffic fatalities were taken from.

I ran two separate regression analyses for each of my four dependent variables; one using the average population density of each city and one using the two population density dummy variables. This should mitigate the effects of any correlation between population density and the dummy variables which represent it. Along with the three measures of population density, each regression also included VMT per capita, two measures of commute mode, and the population of each city's estimated urbanized area as a control variable.

(3.1)

$$Y = \beta_0 + \beta_1 \text{VMT} + \beta_2 \%{>}12{,}500 + \beta_3 \%{<}1{,}500 + \beta_4 \%\text{carpool} + \beta_5 \%\text{don't drive} + \beta_6 \text{population}$$

(3.2)

$$Y = \beta_1 \text{VMT} + \beta_2 \text{density} + \beta_3 \%\text{carpool} + \beta_4 \%\text{don't drive} + \beta_5 \text{population}$$

I ran these regressions using traffic fatalities per 100,000 people, median AQI, percentage of days with 'good' AQI, and average number of vehicles per household as dependent variables.

Chapter 4: Results & Discussion

Every regression model showed significant signs of heteroscedasticity, and as such I employed robust standard errors in each regression.

I: Social Costs

Traffic fatalities per 100,000 people, my proxy for social costs, returned the most significant findings. Both models showed a statistically significant correlation between traffic fatalities per 100,000 people and every independent variable except VMT per capita. The first model, which is recorded in Table 4.1 and used only the population density dummy variables, found that each additional percent of the population living at or above a density of 12,500 people per square mile or greater leads to a decrease of 0.19 traffic fatalities per 100,000 people. Likewise, a 1% increase in the proportion of the population living at or below a density of 1,500 people per square mile leads to an increase of 0.10 traffic fatalities per 100,000 people.

The results relating to measures of commute mode showed that a 1% increase in the proportion of motor vehicle commuters who carpool leads to an increase of 0.75 traffic fatalities per 100,000 people, a finding that contradicts my hypothesis. The model also showed that a 1% increase in the proportion of the population who commutes using a mode of transportation other than a motor vehicle such as a bike, bus, or subway, leads to a decrease of 0.40 traffic fatalities per 100,000 people.

Population, which was used as a control variable, also produced a significant result: An increase in population of about 1.3 million will lead to one fewer traffic death per 100,000 people.

Table 4.1: ***Results for traffic fatalities (density dummies, all cities)***

Traffic fatalities	Coef.	Std. Err.	t	P>t	[95% Conf.	Interval]
VMT per capita	0.0002136	0.000176	1.21	0.228	-0.0001359	0.000563
p>12500	-18.5768	6.779984	-2.74	0.007***	-32.0405	-5.11309
p<1500	9.913407	3.781142	2.62	0.010***	2.404809	17.42201
p_carpool	75.21856	27.76118	2.71	0.008***	20.09034	130.3468
p_other	-40.25561	9.611268	-4.19	0.000***	-59.34168	-21.1695
population	7.73E-07	2.15E-07	3.59	0.001***	3.45E-07	1.20E-06
_cons	6.288044	3.511226	1.79	0.077	-0.6845552	13.26064

*** = significant at a 99% level

The regression model using average population density instead of the density dummy variables is recorded in Table 4.2. Notably, this regression found that an increase in population density of 972.67 people per square mile will lead to one fewer traffic fatality per 100,000 people. VMT per capita, which did not return a significant result in the previous regression, did reach significance here: An increase of 3,338.9 vehicle miles per capita will increase traffic fatalities per 100,000 people by 1. The rest of the results were comparable to those of the previous model.

Table 4.2: ***Results for traffic fatalities (average density, all cities)***

Traffic fatalities	Coef.	Std. Err.	t	P>t	[95% Conf.	Interval]
VMT per capita	0.0002995	0.0001754	1.71	0.091*	-0.0000489	0.0006478
Density	-0.0010281	0.0004675	-2.2	0.030**	-0.0019563	-9.99E-05
p_carpool	64.14798	32.45424	1.98	0.051*	-0.2906691	128.5866
p_other	-52.44305	8.582644	-6.11	0.000***	-69.48409	-35.40201
population	5.73E-07	1.96E-07	2.93	0.004***	1.85E-07	9.62E-07
_cons	12.66986	3.549395	3.57	0.001	5.622457	19.71727

* = Significant at a 90% level
** = Significant at a 95% level
*** = Significant at a 99% level

Traffic fatalities per 100,000 people range from 3.97 to 26.26 (See Figure 4.1). The proportion of the population living at a density of 12,500 or greater ranges from 0 to 56% (See Figure 4.2), and the proportion of the population living at a density of 1,500 people per square

mile or less ranges from 2% to 49% (See Figure 4.3). This should put into perspective the significance of the effect of density and transit mode on social costs related to automobile dependency. In the average city, about 21% of the population lives at a density of 1,500 people per square mile or lower and about 5% lives at a density of 12,500 people per square mile or higher, enough to account for an increase of 2.05 or a decrease of 0.98 traffic fatalities, respectively. I would argue that these are both economically significant results. Despite the fact that less than half of the MSAs sampled have a population greater than 1.3 million, I would argue that even the population control variable produces an economically significant result. In New York City, for instance, population, which is around 20 million, is high enough to account for a decrease of more than 15 fatalities per 100,000 people. For reference, New York City only had 4.6 traffic fatalities per 100,000 people in 2021.

Figure 4.1: ***Distribution of Traffic Fatalities per 100,000 People***

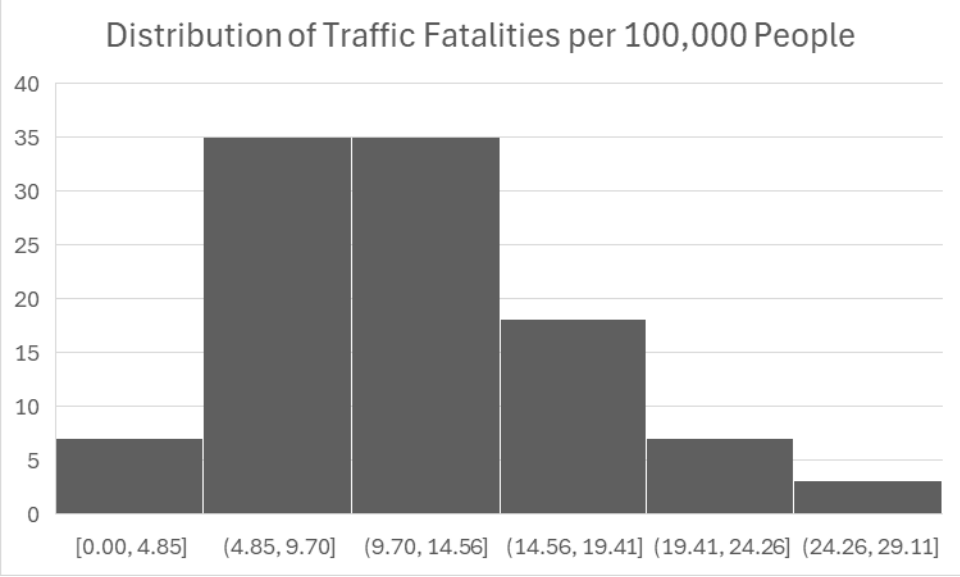


Figure 4.2: ***Distribution of P>12,500***

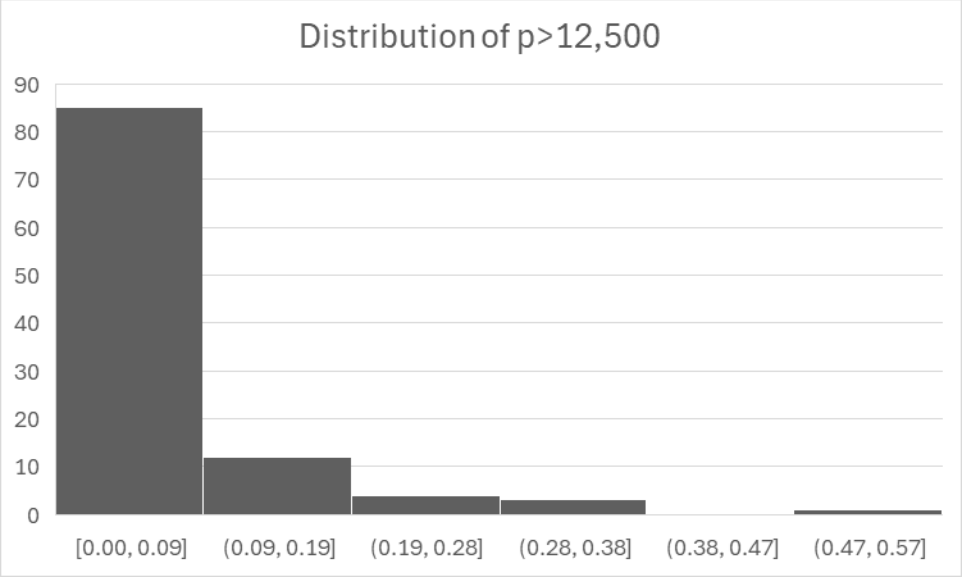
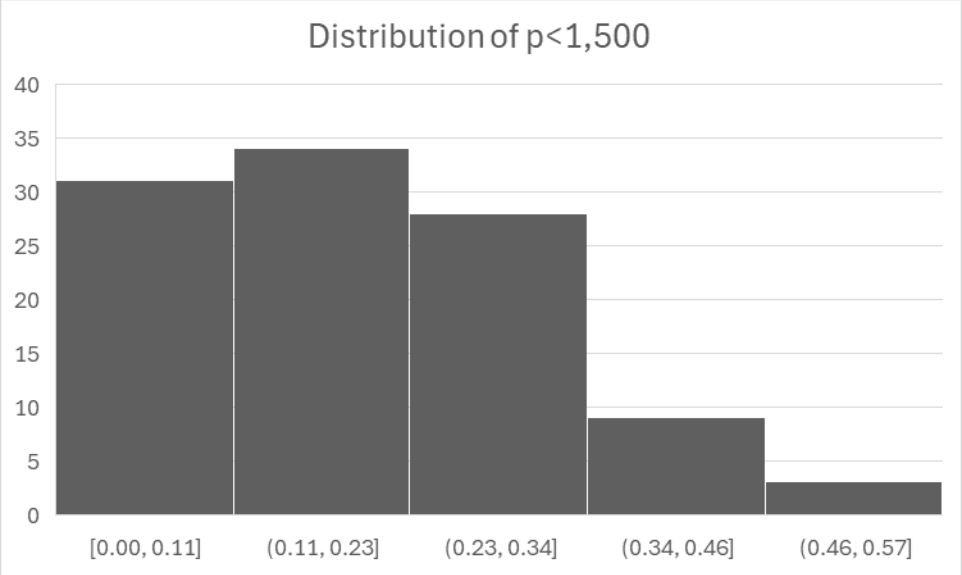
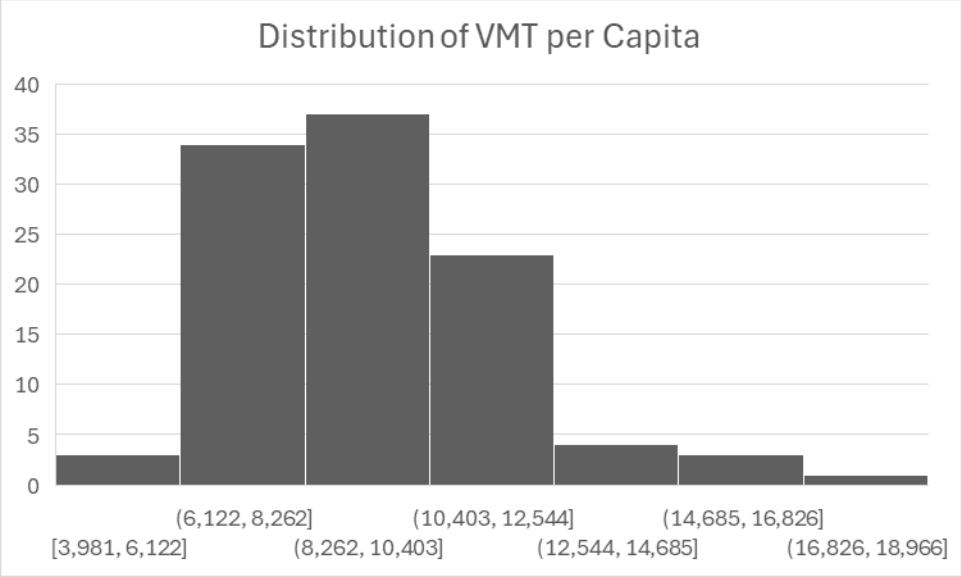


Figure 4.3: ***Distribution of P<1,500***



VMT per capita ranges from 5,202.21 to 18,533.71, with an average of 9,581.49 miles per capita (See Figure 4.4). This means that VMT is enough to account for at least one traffic fatality per 100,000 people in every city sampled, and around three fatalities on average.

Figure 4.4: ***Distribution of Vehicle Miles Travelled per Capita***



The most surprising outcome of this regression was the positive correlation between the number of carpoolers and the number of traffic fatalities. My initial reasoning was that a city with a high proportion of carpoolers is one that is less automobile dependent and therefore should exhibit lower social costs associated with automobile dependency. I suspect this result is actually due to the fact that more carpoolers means there will be more people in the average car, which means more people will be seriously injured or killed in automobile accidents.

I was also surprised that VMT per capita was not significant in both regressions. It could be that the values for VMT per capita are very close to one another, as in Figure 4.4, making a meaningful correlation difficult to find. It could also be that VMT per capita is highly correlated with population density or commute mode, which would make sense and would interfere with my results. Figures 4.5 and 4.6 below, however, show that the correlation between average population density and VMT per capita and the correlation between the proportion of commuters who do not use a motor vehicle and VMT per capita are fairly weak. There does appear to be

some negative correlation, which I would expect because each of these variables are proxies for automobile dependency, but not enough to meaningfully interfere with results.

Figure 4.5: ***Correlation Between VMT and Average Density***

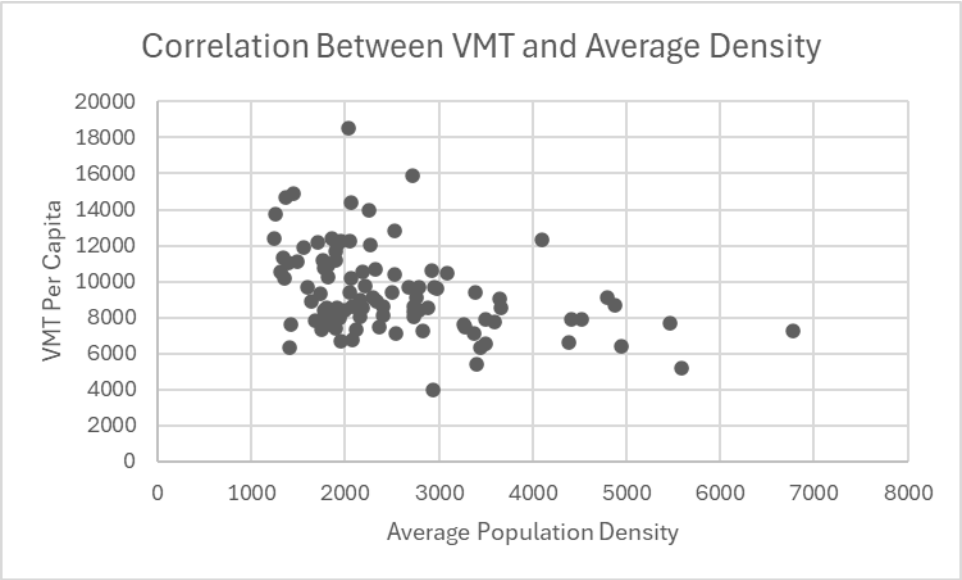
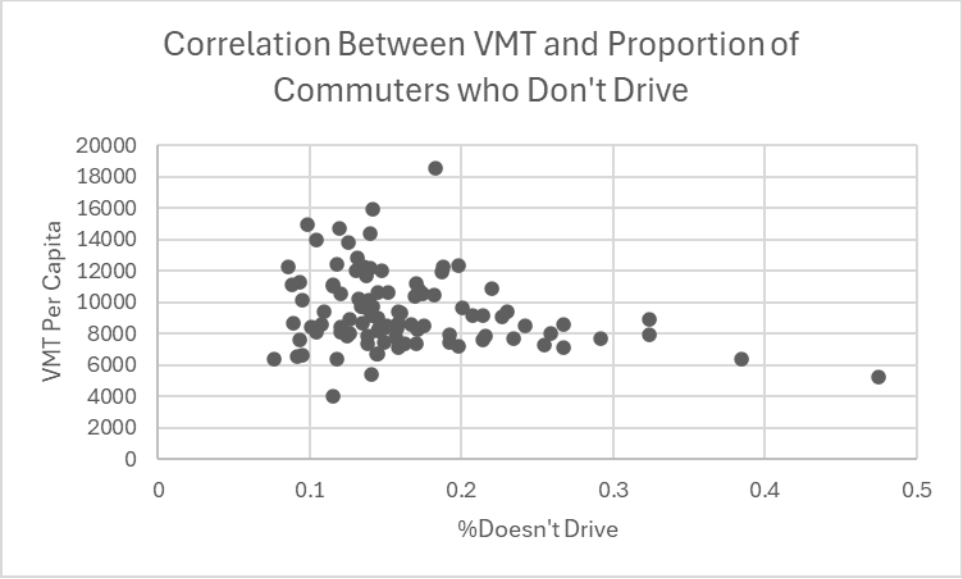


Figure 4.6: ***Correlation Between VMT and Proportion of Commuters who Don't Drive***



My data collection may have also been the issue. Data for VMT per capita came from each city's actual urbanized area, whereas data for most of the other variables was taken from my estimated urbanized area. In many cases, the difference in area and population between the two

was larger than I had expected. To control for this, I ran the same regressions again but only including cities where the population of the estimated urbanized area was within 25% of the population of the actual urbanized area. This narrowed the sample size from 100 cities to 81.

Both regressions run this way are displayed in tables 4.3 and 4.4. Each found a significant correlation between traffic fatalities per 100,000 people and VMT per capita with a similar coefficient to the previous regression despite the sample size being smaller than before. This suggests that the discrepancy before, whereby one regression returned a significant finding while the other did not, is due to faults in my data collection methods rather than a dubious correlation.

The other notable difference is that the regression using average density as an independent variable rather than the population density dummies did not return a significant correlation between the proportion of motor vehicle commuters who carpool and traffic fatalities.

Table 4.3: ***Results for traffic fatalities (density dummies, filtered cities)***

Traffic fatalities	Coef.	Std. Err.	t	P>t	[95% Conf.	Interval]
VMT per capita	0.000396	0.000193	2.05	0.044**	1.09E-05	0.000782
p>12500	-21.9377	7.37797	-2.97	0.004***	-36.6387	-7.2368
p<1500	12.11645	3.661068	3.31	0.001***	4.821616	19.41129
p_carpool	46.63038	25.22496	1.85	0.069*	-3.63145	96.89221
p_other	-34.8945	9.51166	-3.67	0.000***	-53.8469	-15.9421
population	9.79E-07	2.43E-07	4.04	0.000***	4.96E-07	1.46E-06
_cons	5.624457	3.32798	1.69	0.095	-1.00669	12.2556

* = Significant at a 90% level
** = Significant at a 95% level
*** = Significant at a 99% level

Table 4.4: ***Results for traffic fatalities (average density, filtered cities)***

Traffic fatalities	Coef.	Std. Err.	t	P>t	[95% Conf.	Interval]
VMT per capita	0.0004671	0.0001974	2.37	0.021**	0.0000739	0.00086
Density	-0.001487	0.0004824	-3.08	0.003***	-0.002448	-0.000526
p_carpool	39.93159	28.07264	1.42	0.159	-15.99197	95.85515
p_other	-48.08874	8.966051	-5.36	0.000***	-65.95003	-30.22745
population	7.70E-07	2.13E-07	3.62	0.001***	3.46E-07	1.19E-06
_cons	13.23872	3.245984	4.08	0.000	6.772384	19.70505

** = Significant at a 95% level
*** = Significant at a 99% level

II: Environmental Costs

I used two measures of air quality as proxies for environmental costs of automobile dependency: Median AQI and the percentage of days with an AQI below 50, or "good". The regressions using median AQI as the dependent variable are recorded in tables 4.5 and 4.6. Both reported statistically significant correlations between median AQI and one of the commute mode variables, and well as with the population control variable. A 1% increase in commuters who use any mode of transportation other than a motor vehicle leads to a decrease in AQI by 0.48 (Table 4.5) or 0.77 (Table 4.6) points. These findings are consistent with my hypothesis as a lower AQI score represents better air quality. Additionally, an increase in population of 379,000 (Table 4.5) or 521,000 (Table 4.6) leads to an increase in AQI of one point.

The regression using the density dummies, displayed in Table 4.5, also found that a 1% increase in the proportion of the population living at a density of 1,500 people per square mile or lower leads to a decrease in AQI of 0.20 points. This finding is not in accordance with my hypothesis.

Table 4.5: ***Results for median AQI (density dummies)***

Median AQI	Coef.	Std. Err.	t	P>t	[95% Conf.	Interval]
VMT per capita	-0.0003289	0.000485	-0.68	0.499	-0.0012912	0.000633
p>12500	-44.12528	28.46224	-1.55	0.124	-100.6457	12.3951
p<1500	-20.29293	8.318913	-2.44	0.017**	-36.81264	-3.77322
p_carpool	110.6815	105.6095	1.05	0.297	-99.038	320.401
p_other	-48.26625	22.551	-2.14	0.035**	-93.04806	-3.48444
population	2.64E-06	9.07E-07	2.91	0.005***	8.35E-07	4.44E-06
_cons	53.96401	12.33722	4.37	0.000	29.46473	78.46329

** = Significant at a 95% level
*** = Significant at a 99% level

Table 4.6: ***Results for median AQI (average density)***

Median AQI	Coef.	Std. Err.	t	P>t	[95% Conf.	Interval]
VMT per capita	-0.0001266	0.0005231	-0.24	0.809	-0.0011652	0.0009121
Density	0.002002	0.0015626	1.28	0.203	-0.0011006	0.0051046
p_carpool	75.97373	139.994	0.54	0.589	-201.9876	353.9351
p_other	-77.40993	22.31623	-3.47	0.001***	-121.7193	-33.10053
population	1.92E-06	8.14E-07	2.36	0.020**	3.03E-07	3.53E-06
_cons	50.23298	13.10958	3.83	0.000	24.2036	76.26237

** = Significant at a 95% level
*** = Significant at a 99% level

Regressions using the percentage of days with a recorded AQI of 50 or lower returned significant findings for two of the three measures of population density as well as the proportion of commuters who use any mode of transport other than a motor vehicle and the population control variable, the same as in the previous two regressions. The regression employing the two population density dummy variables, displayed in Table 4.7, found that a 1% increase in the proportion of the population living at a density of 1,500 people per square mile or lower increases the proportion of days with good AQI by 0.39%. The regression employing average population density as an independent variable, displayed in Table 4.8, found that an increase in density of 192.68 people per square mile decreases the proportion of days with good AQI by 1%. Similarly to the findings with median AQI, both of these findings contradict my hypothesis.

Table 4.7: ***Results for percentage of days with AQI<50 (density dummies)***

p_Good AQI	Coef.	Std. Err.	t	P>t	[95% Conf.	Interval]
VMT per capita	2.67E-06	8.14E-06	0.33	0.744	-0.0000135	1.88E-05
p>12500	0.5403334	0.391706	1.38	0.171	-0.2375178	1.318185
p<1500	0.3919143	0.178964	2.19	0.031**	0.0365283	0.7473
p_carpool	0.1956242	1.445382	0.14	0.893	-2.674618	3.065866
p_other	0.881769	0.361623	2.44	0.017**	0.1636571	1.599881
population	-4.54E-08	1.05E-08	-4.31	0.000***	-6.63E-08	-2.45E-08
_cons	0.2308695	0.18342	1.26	0.211	-0.133367	0.595106

** = Significant at a 95% level
*** = Significant at a 99% level

Table 4.8: ***Results for percentage of days with AQI<50 (average density)***

p_Good AQI	Coef.	Std. Err.	t	P>t	[95% Conf.	Interval]
VMT per capita	-7.10E-07	8.07E-06	-0.09	0.930	-0.0000167	0.0000153
Density	-0.0000519	0.0000243	-2.14	0.035**	-0.0001002	-3.67E-06
p_carpool	1.035585	1.915414	0.54	0.590	-2.767515	4.838685
p_other	1.242753	0.3090711	4.02	0.000***	0.6290854	1.856421
population	-3.34E-08	9.49E-09	-3.52	0.001***	-5.22E-08	-1.46E-08
_cons	0.3337456	0.2017936	1.65	0.101	-0.0669204	0.7344116

** = Significant at a 95% level
*** = Significant at a 99% level

These results are all economically significant, including the correlation between both measures of environmental sustainability and the population control variable, although the implications of this are limited. I was surprised by the finding that more people living at a lower density leads to higher air quality. My thinking was that lower density correlates with more urban sprawl and therefore higher levels of driving and more pollution. However, it makes sense that denser areas would have poorer air quality as there are more people there to pollute them. On the other hand, rural areas would have better air quality because nobody is living there. This result seems to imply that higher density areas are worse for the environment, which I doubt is the case. Rather, people in urban areas probably drive less and pollute less, but the pollution that exists is concentrated in one area and leads to lower air quality. I suspect that this result would be different if I had used a measure that controlled for population, such as carbon dioxide emissions per capita, which is what was used by Zheng et al. (2011) and Garceau et al. (2013). This is a potential area for future research.

III: Economic Costs

I employed the average number of vehicles per household across estimated urbanized areas as a measure of economic costs. This is consistent with previous research, which used the number of vehicles per household combined with the estimated cost of vehicle ownership to

estimate the amount each household spent on transportation (Zheng et al., 2011, Garceau et al., 2013). The estimated cost of vehicle ownership is constant across geographies, so it will not change from city to city. As such, I simply used the number of vehicles per household on its own to simplify things.

The regression model returned significant correlations between the average number of vehicles per household and all measures of population density. As predicted, the correlation between the average number of vehicles per household was negatively correlated with the proportion of the population living at a density of 12,500 or greater: a 1% increase in the proportion of the population living at a density of 12,500 or greater led to a decrease of 0.005 vehicles per household. However, a 1% increase in the proportion of the population living at a density of 1,500 or lower also corresponded to a decrease of 0.003 vehicles per household, a finding that is not in line with my prediction. These results can be seen in Table 4.9.

The regression using the population density dummy variables also found that a 1% increase in the proportion of motor vehicle commuters who carpool increased the average number of vehicles per household by 0.03.

Table 4.9: ***Results for vehicles per household (density dummies, all cities)***

Vehicles per household	Coef.	Std. Err.	t	P>t	[95% Conf.	Interval]
VMT per capita	2.76E-06	6.47E-06	0.43	0.670	-0.0000101	1.56E-05
p>12500	-0.5037539	0.244284	-2.06	0.042**	-0.9888526	-0.01866
p<1500	-0.3843168	0.144708	-2.66	0.009***	-0.6716788	-0.09695
p_carpool	2.929364	1.02444	2.86	0.005***	0.8950302	4.963698
p_other	-0.378036	0.321069	-1.18	0.242	-1.015615	0.259543
population	-5.19E-09	8.56E-09	-0.61	0.546	-2.22E-08	1.18E-08
_cons	1.603609	0.141762	11.31	0.000	1.322097	1.885121

** = Significant at a 95% level
*** = Significant at a 99% level

The regression using average density as an independent variable rather than the population density dummies found a significant correlation between the average number of vehicles per household and the average population density of a city as well as the proportion of commuters who use any mode of transportation other than a motor vehicle and the population control variable. The model shows that that an increase in density of 18,050.54 people per square mile corresponds to an increase of one vehicle per household. This result is contradictory to my hypothesis.

As predicted, a 1% increase in the proportion of commuters who use any mode of transportation other than a motor vehicle leads to a decrease of 0.007 vehicles per household. The population control variable showed that an increase in population of 57.5 million leads to a decrease of 1 vehicle per household. These results are illustrated in Table 4.10.

Table 4.10: ***Results for vehicles per household (average density, all cities)***

Vehicles per household	Coef.	Std. Err.	t	P>t	[95% Conf.	Interval]
VMT per capita	6.27E-06	6.96E-06	0.9	0.369	-7.54E-06	0.0000201
Density	0.0000554	0.0000223	2.49	0.015**	0.0000112	0.0000997
p_carpool	1.988561	1.306659	1.52	0.131	-0.6058407	4.582963
p_other	-0.7160463	0.2487779	-2.88	0.005***	-1.210001	-0.222092
population	-1.74E-08	8.84E-09	-1.97	0.051*	-3.50E-08	1.11E-10
_cons	1.501372	0.1321716	11.36	0	1.238942	1.763802

* = Significant at a 90% level
** = Significant at a 95% level
*** = Significant at a 99% level

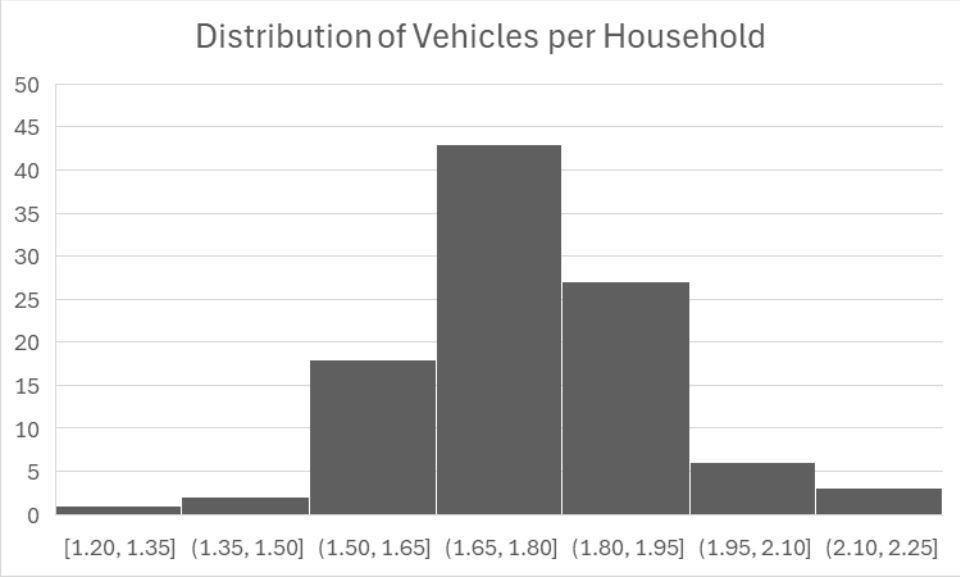
Once again, the lack of significant results for VMT per capita was unexpected. Similarly to my regression of social costs, I tried to account for the shortcomings of my data collection methods by running a regression that included only the 81 selected estimated urbanized areas whose populations were within 25% of their respective actual urbanized areas. These regressions both returned significant correlations between the average number of vehicles per household and

VMT per capita. The rest of the results were comparable to those of the regressions using the entire sample.

The regression employing the population density dummy variables found that an increase of 95,208 vehicle miles per capita corresponds to an increase of 1 vehicle per household. The regression employing the average population density of each city found that an increase of 177,000 vehicle miles per capita corresponds to an increase of 1 vehicle per household. The unexpected results showing a positive correlation between population density and the average number of vehicles per household and the negative correlation between the proportion of the population living at a density of 1,500 people per square mile or less and the average number of vehicles per household remained the same.

The results of this regression are all economically significant. Even though some of the numbers required to change the average number of vehicles per household by just 1 is quite large (95,000 or 177,000 in the case of VMT per capita and 57.5 million in the case of population density), an increase of just a fraction of a vehicle is still a large change: The average number of vehicles per household across all cities only ranges from 1.20 to 2.22. This distribution is shown in Figure 4.7.

Figure 4.7: Distribution of Vehicles per Household

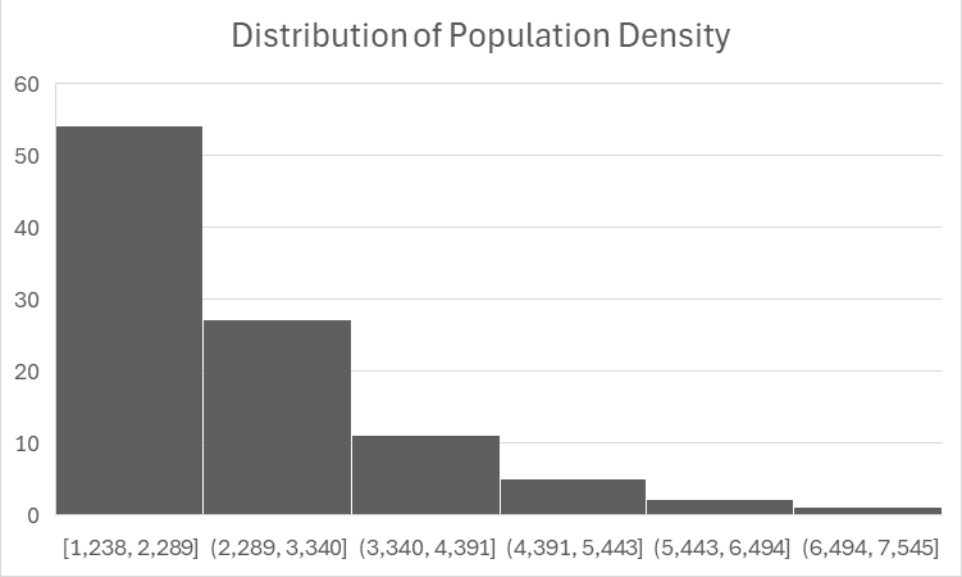


The most surprising findings of this research are the positive correlation between the average number of vehicles per household and the average population density of a city, and the negative correlation between the proportion of the population living at a density of 1,500 people per square mile or less and the average number of vehicles per household. Several observations lead me to believe that this is not a reliable conclusion. Firstly, the proportion of the population living at a density of 12,500 or greater was negatively correlated with the average number of vehicles per household, which directly contradicts the finding that higher average density increases it. Secondly, my observations during data collection showed that the average number of vehicles per household were lower across tracts with higher densities and higher across tracts with lower densities. A study using individual tracts as observations rather than entire urbanized areas would likely show a different result and may be an interesting topic to be explored by future research.

My last, though hardly scientific, observation is that many cities that I would expect to show signs of urban sprawl tended to have higher densities than some that I wouldn't. For

instance, New York City had an average population density of 5,592.95 people per square mile, whereas Los Angeles had an average population density of 6,778.33 people per square mile, despite New York City being famously dense and walkable and Los Angeles being infamously the opposite. Other cities such as Stockton and Oxnard, California had densities of 4,382.97 and 3,406.80 people per square mile, respectively, very high compared to most cities of a similar size, and an average of 2.05 and 2.08 vehicles per household, respectively, some of the highest of any city sampled. For these reasons, I think that one or more of my density variables are somehow not accurate due to my data collection methods, or that one or more of my density variables are simply not good measures of urban sprawl.

Figure 4.8: Distribution of Population Density



IV: Vehicles per Household as an Independent Variable

Using the average number of vehicles per household as a measure of the economic cost of automobile dependency presents a few issues. One is that it doesn't scale based on use. A household that lives in a car dependent city and must drive farther and more often could have the

same number of vehicles as a household in a walkable, dense environment that rarely drives, but the amount that each household spends on gas and maintenance for their vehicles will be vastly different. Likewise, the family in the walkable city might use other forms of transportation like taxi, bus, or subway whose cost is not factored in when only their private motor vehicles are considered.

Additionally, some measures of automobile dependency are not completely independent of the average number of vehicles per household. For example, if the average number of vehicles per household is higher, it would make sense that VMT per capita would be higher because there are more drivers on the road. Even population density could be affected by the average number of vehicles per household due to the space they take up and because of the "feedback loop" described by Silva and Altieri (2004). For these reasons, it makes more sense to use the average number of vehicles per household as a measure of automobile dependency rather than an effect of it. I was prevented from doing this for economic costs because I couldn't find a better dataset, but I did try running the social and environmental regressions again employing the average number of vehicles per household as independent variables. The outcome was significant in the two environmental regressions, which showed that one additional vehicle per household increased a city's median AQI by 20.37 points (Table 4.11) or 25.23 points (Table 4.12) and decreased the proportion of days with an AQI below 50 by 37.17% (Table 4.13) or 41.55% (Table 4.14). These results are both economically significant and in agreement with my hypothesis; a city in which the average household has more vehicles is one that should be more automobile dependent and thus should exhibit higher sustainability costs associated with automobile dependency.

Table 4.11: ***Results for median AQI with vehicles per household as an IV (density dummies)***

Median AQI	Coef.	Std. Err.	t	P>t	[95% Conf.	Interval]
VMT per capita	-0.00039	0.000463	-0.83	0.407	-0.0013	0.000534
p>12500	-33.8662	27.62992	-1.23	0.223	-88.7416	21.00921
p<1500	-12.4662	7.798285	-1.6	0.113	-27.9543	3.021847
p_carpool	51.02426	105.7206	0.48	0.631	-158.946	260.9945
p_other	-40.5675	21.24363	-1.91	0.059*	-82.7591	1.624233
population	2.74E-06	8.58E-07	3.2	0.002***	1.04E-06	4.45E-06
Vehicles per household	20.36525	7.208177	2.83	0.006***	6.049192	34.68131
_cons	21.3061	14.62506	1.46	0.149	-7.74052	50.35273

* = Significant at a 90% level

*** = Significant at a 99% level

Table 4.12: ***Results for median AQI with vehicles per household as an IV (average density)***

Median AQI	Coef.	Std. Err.	t	P>t	[95% Conf.	Interval]
VMT per capita	-0.00028	0.000489	-0.58	0.562	-0.00126	0.000686
Density	0.000604	0.001628	0.37	0.712	-0.00263	0.003836
p_carpool	25.80471	125.2823	0.21	0.837	-222.981	274.5905
p_other	-59.3449	18.96287	-3.13	0.002***	-97.0014	-21.6884
population	2.36E-06	8.03E-07	2.94	0.004***	7.64E-07	3.95E-06
Vehicles per household	25.22881	8.621878	2.93	0.004***	8.107466	42.35015
_cons	12.35517	16.49478	0.75	0.456	-20.4002	45.11054

*** = Significant at a 99% level

Table 4.13: ***Results for percentage of days with AQI<50 with vehicles per household as an IV (density dummies)***

p_Good AQI	Coef.	Std. Err.	t	P>t	[95% Conf.	Interval]
VMT per capita	3.70E-06	8.15E-06	0.45	0.651	-1.3E-05	1.99E-05
p>12500	0.35307	0.371561	0.95	0.344	-0.38488	1.091023
p<1500	0.24905	0.163065	1.53	0.130	-0.07481	0.572911
p_carpool	1.284573	1.462718	0.88	0.382	-1.62051	4.189657
p_other	0.74124	0.361813	2.05	0.043**	0.022649	1.45983
population	-4.73E-08	9.40E-09	-5.04	0.000***	-6.60E-08	-2.87E-08
Vehicles per household	-0.37174	0.113019	-3.29	0.001***	-0.5962	-0.14727
_cons	0.826988	0.215531	3.84	0.000	0.398925	1.255051

** = Significant at a 95% level

*** = Significant at a 99% level

Table 4.14: ***Results for percentage of days with AQI<50 with vehicles per household as an IV (average density)***

p_Good AQI	Coef.	Std. Err.	t	P>t	[95% Conf.	Interval]
VMT per capita	1.90E-06	8.06E-06	0.24	0.815	-1.4E-05	1.79E-05
Density	-2.9E-05	2.28E-05	-1.27	0.209	-7.4E-05	1.65E-05
p_carpool	1.861855	1.75558	1.06	0.292	-1.62438	5.34809
p_other	0.945228	0.301066	3.14	0.002***	0.34737	1.543086
population	-4.07E-08	8.58E-09	-4.74	0.000***	-5.77E-08	-2.36E-08
Vehicles per household	-0.41551	0.120857	-3.44	0.001***	-0.65551	-0.17551
_cons	0.957583	0.19927	4.81	0	0.561871	1.353294

*** = Significant at a 99% level

Chapter 5: Conclusion

I predicted that cities with higher measures of automobile dependency would exhibit worse sustainability outcomes in social, environmental, and economic categories. I expected VMT per capita to be positively associated with traffic fatalities and the number of vehicles per household and negatively associated with measures of air quality, meaning an increase in VMT would lead to worse air quality, not lower AQI (lower AQI is better). I expected measures of population density to be negatively associated with traffic fatalities and vehicles per household and positively associated with measures of air quality. I expected commuters carpooling or taking other means of transport to work such as buses, subways, or bikes to be negatively associated with traffic fatalities and vehicles per household and positively associated with measures of air quality.

My findings showed that VMT per capita is positively associated with traffic fatalities and vehicles per household. Density is negatively associated with traffic fatalities. The proportion of commuters using any mode of transportation other than a motor vehicle is negatively associated with traffic fatalities and vehicles per household and positively associated with measures of air quality. These results all align with my hypothesis. I also found that the average number of vehicles per household is positively associated with measures of air quality.

My research returned a number of unexpected results. The proportion of drivers who carpool was associated with an increase in traffic fatalities. This makes sense because if more people are carpooling, there will be more people in the average car, so car accidents should be more deadly than they would be otherwise. I expected that a city where more people carpool would be one that is less automobile dependent and would therefore exhibit lower social costs from driving.

Density was associated with poorer air quality. I expected that higher density would lead to less pollution, and therefore better air quality. However, while there is likely less pollution per person in dense areas, the amount is concentrated, which means the air quality is worse than less dense areas.

VMT per capita, the most straightforward measure of automobile dependency, surprisingly failed to return to significant results in some regressions. This could be due to correlation between VMT per capita and the other variables used in my regressions such as population density and commute mode, although this correlation appears weak. It could also be that VMT values across cities are very similar to one another, making a meaningful correlation difficult to find. I think the most likely issue was with my methodology and the difference in areas that data for different variables was collected from. When I tried to control for error in my methodology, VMT per capita returned significant results.

The most surprising result was a positive correlation between density and the number of vehicles per household. This result is not easy to explain. It may be due to inconsistencies in my data collection methods, or due to population density being an insufficient proxy for urban sprawl. I seriously doubt that denser cities are ones with higher vehicle ownership rates as this result seems to suggest. This is because one of the density variables contradicted the other two in regressions for vehicle ownership, and because my observations while collecting data suggest that households in denser areas of cities have fewer cars than those in suburban areas.

Potential areas for future study might address some of these unexpected findings. The most interesting thing to explore would be the correlation between density and vehicle ownership. Looking at these variables on a smaller scale than I did in this paper, at the census tract level, for instance, would likely yield a result that aligns more with my hypothesis than my

findings did. It would also be interesting to explore why some cities that seem representative of urban sprawl in most other regards have higher than expected densities. The correlation between density and air quality would also make for interesting research. I expect that a study using per capita data, like carbon dioxide emissions per capita, rather than air quality across the whole city would yield results that align more with my hypotheses than my findings did. The implications of the finding that carpooling leads to more traffic fatalities could also be explored further.

References

Ahangari, H., Atkinson-Palombo, C., & Garrick, N. W. (2017). Automobile-dependency as a barrier to vision zero, evidence from the states in the USA. *Accident Analysis and Prevention, 107*, 77–85. 10.1016/j.aap.2017.07.012

Bureau of Transportation Statistics.*Vehicle Ownership and Commuting Most Recent (2021) Selected Variables. .* https://equity-data.dot.gov/datasets/usdot::vehicle-ownership-and-commuting-most-recent-2021-selected-variables/about

Cervero, R., & Murakami, J. (2010). Effects of Built Environments on Vehicle Miles Traveled: Evidence from 370 US Urbanized Areas. *Environment and Planning.A, 42*(2), 400–418. 10.1068/a4236

Edlin, A. S., & Karaca-Mandic, P. (2006). The Accident Externality from Driving. *The Journal of Political Economy, 114*(5), 931–955. 10.1086/508030

Ewing, R., Pendall, R., Chen, D., & TRB. (2003). Measuring sprawl and its transportation impacts: PLANNING AND ADMINISTRATION. *Transportation Research Record*, (1831), 175–183. 10.3141/1831-20

Frank, L. D., Andresen, M. A., & Schmid, T. L. (2004). Obesity relationships with community design, physical activity, and time spent in cars. *American Journal of Preventive Medicine, 27*(2), 87–96. 10.1016/j.amepre.2004.04.011

Frederick, C., Riggs, W., & Gilderbloom, J. H. (2018). Commute mode diversity and public health: A multivariate analysis of 148 US cities. *International Journal of Sustainable Transportation, 12*(1), 1–11. 10.1080/15568318.2017.1321705

Garceau, T., Atkinson-Palombo, C., Garrick, N., Outlaw, J., McCahill, C., & Ahangari, H. (2013). Evaluating selected costs of automobile-oriented transportation systems from a sustainability perspective. *Research in Transportation Business & Management, 7*, 43–53. 10.1016/j.rtbm.2013.02.002

Kang, S. (2021). Beyond households: regional determinants of housing instability among low-income renters in the United States. *Housing Studies*, *36*(1), 80–109. 10.1080/02673037.2019.1676402

Kooshian, C., & Winkelman, S. (2011). *Growing Wealthier: Smart Growth, Climate Change and Prosperity*. . http://growingwealthier.info/docs/growing_wealthier.pdf

Lu, P., Hamori, S., Sun, L., & Tian, S. (2024). Does the electric vehicle industry help achieve sustainable development goals?—evidence from China. *Frontiers in Environmental Science*, 10.3389/fenvs.2023.1276382

National Highway Traffic Safety Administraion. (2024). *Fatality Analysis Reporting System*. . https://www.nhtsa.gov/file-downloads?p=nhtsa/downloads/FARS/2021/National/

Noland, R. B., & Lem, L. L. (2002). A review of the evidence for induced travel and changes in transportation and environmental policy in the US and the UK. *Transportation Research.Part D, Transport and Environment, 7*(1), 1–26. 10.1016/S1361-9209(01)00009-8

Office of Highway Policy Information. (2022). *Highway Statistics 2022*. . https://www.fhwa.dot.gov/policyinformation/statistics/2022/

Office of Highway Policy Information. (2023). *Highway Statistics 2021 Table HM-71*. . https://www.fhwa.dot.gov/policyinformation/statistics/2021/hm71.cfm

Parry, I. W. H., Walls, M., & Harrington, W. (2007). Automobile Externalities and Policies. *Journal of Economic Literature, 45*(2)10.1257/jel.45.2.373

Shirley, C., & Winston, C. (2004). Firm inventory behavior and the returns from highway infrastructure investments. *Journal of Urban Economics, 55*(2), 398–415. 10.1016/j.jue.2003.11.001

Silva, C., & Altieri, M. (2022). Is regional accessibility undermining local accessibility? *Journal of Transport Geography, 101*, 103336. 10.1016/j.jtrangeo.2022.103336

Stone, B. (2008). Urban sprawl and air quality in large US cities. *Journal of Environmental Management, 86*(4), 688–698. 10.1016/j.jenvman.2006.12.034

United States Census Bureau. (2020a). *Housing Units.* https://data.census.gov/table/ACSDT5Y2020.B25001?q=United+States&t=Housing+Units&g=010XX00US,$1400000&y=2020

United States Census Bureau. (2020b). *Total Population.* . https://data.census.gov/table/DECENNIALDHC2020.P1?q=P1&g=010XX00US,$1400000&d=DEC%20Demographic%20and%20Housing%20Characteristics

United States Census Bureau. (2021). *TIGER Shapefiles.* . https://www2.census.gov/geo/tiger/TIGER2021/TRACT/

United States Census Bureau. (2022, Dec 22). *Redefining Urban Areas following the 2020 Census.* https://www.census.gov/newsroom/blogs/random-samplings/2022/12/redefining-urban-areas-following-2020-census.html

United States Environmental Protection Agency. (2021). *Air Quality Statistics Report.* . https://www.epa.gov/outdoor-air-quality-data/air-quality-statistics-report

White, M. J. (2004). The "Arms Race" on American Roads: The Effect of Sport Utility Vehicles and Pickup Trucks on Traffic Safety. *The Journal of Law & Economics, 47*(2), 333–355. 10.1086/422979

Wolman, H., Galster, G., Hanson, R., Ratcliffe, M., Furdell, K., & Sarzynski, A. (2005). The Fundamental Challenge in Measuring Sprawl: Which Land Should Be Considered? *The Professional Geographer, 57*(1), 94–105. 10.1111/j.0033-0124.2005.00462.x

Zheng, J., Atkinson-Palombo, C., McCahill, C., O'Hara, R., & Garrick, N. W. (2011). Quantifying the Economic Domain of Transportation Sustainability. *Transportation Research Record, 2242*(1), 19–28. 10.3141/2242-03

Zhu, P., Tan, X., Zhao, S., Shi, S., & Wang, M. (2022). Land use regulations, transit investment, and commuting preferences. *Land use Policy, 122*, 106343. 10.1016/j.landusepol.2022.106343