SPATIAL ANALYSIS OF URBAN BUILT ENVIRONMENTS

AND VEHICLE TRANSIT BEHAVIOR

by

Daniel Currie Eisman

A Thesis Presented to the
FACULTY OF THE USC GRADUATE SCHOOL
UNIVERSITY OF SOUTHERN CALIFORNIA
In Partial Fulfillment of the
Requirements for the Degree
MASTER OF SCIENCE
(GEOGRAPHIC INFORMATION SCIENCE AND TECHNOLOGY)

August 2012

Acknowledgments

I would like to thank my Committee Chair, Dr. Robert Vos, whose feedback and support throughout this process have been invaluable. I would also like to thank my other committee members, Dr. Katsuhiko Oda and Dr. Daniel Warshawsky, for their time and effort in assisting me. I am also grateful to Dr. Meredith Franklin for her help with the processing and interpretation of data.

Table of Contents

List of Tables

List of Figures

Abstract

In an effort to explore smart growth principles, this study offers an empirical test of the influence of the built environment at the neighborhood scale on vehicle transit behavior. Using U.S. Census data combined with spatial analysis techniques, the study conducts a cross-sectional analysis of the effect of the built environment on household automobile ownership and vehicles miles traveled (VMTs) in 75 block groups across five metropolitan statistical areas. Variables are measured for density, job and retail access, transit accessibility, and street connectivity. The study also considers confounding variables including household income, regional density, extent of regional transit network, age of neighborhood population, and individual transit expenditure. From these data, best-fit regression models are developed for VMTs and automobile ownership. Although there is significant unexplained variation, the regression models confirm a statistically significant association of VMTs and automobile ownership with the built environment. Among the implications of these findings are that (1) neighborhood density should be encouraged in areas well-served by transit, (2) transit and smart-growth projects will have a greater impact on VMTs in regions that have robust, existing transit systems, and (3) new transit projects will likely be most effective in reducing vehicle ownership if planners focus on better serving moderate and low-income neighborhoods. Future research should examine statistical associations longitudinally, based on updated data from the 2010 U.S. Census, and should attempt to gather primary data on VMTs at the household and neighborhood scales.

Chapter 1: Introduction

This study assesses the effect of the built environment on transit behavior. Specifically, the focus is on travel behavior in terms of vehicle transit by measuring vehicle miles traveled (VMTs) and automobile ownership. While there are other aspects to transit behavior, such as transit ridership, biking, and walking, this study focuses on VMTs and automobile ownership because these variables best capture vehicle behavior.

The purpose of this study is to put the transit-oriented development principles of smart growth to a robust empirical test by exploring the relationship between the built environment and VMTs and automobile ownership. If built environments have an influence on transportation behavior, it is expected that people who live in neighborhoods with built environments that include features of transit-oriented development will drive less than those who live in neighborhoods more characteristic of sprawl morphologies.

This is accomplished by assessing a set of built environment factors at the neighborhood and regional scales. The following independent variables are measured in this analysis: neighborhood density, job access, retail access, transit access, and street connectivity. The confounding variables measured include household income, regional density, extent of regional transit network, age of neighborhood population, and individual transit expenditure.

Using a cross-sectional study design, this research seeks to find the spatial relationships and variables that drive automobile ownership levels and VMTs at the neighborhood and regional scales. Variables are measured for 75 sample block groups across five metropolitan statistical areas (MSAs)—Chicago, IL; Miami, FL; Portland,

OR; San Diego, CA; and Washington, DC. The findings confirm a significant association with some smart growth principles and lower automobile ownership and VMTs.

1.1 Motivation

The sprawl style suburban development that became commonplace in the post-World War II years has been blamed for several negative societal and economic impacts. Traffic gridlock has been linked to the proliferation of far-flung, sprawling automobile-dependent suburbs (Gordon & Richardson, 1997). Road networks lack adequate capacity to keep up with the constantly growing demand. Even if government agencies had adequate funding to build new roads, many see new road construction as an inefficient and unsustainable way of meeting the transportation demands of the public. Gridlock has both societal and economic costs in that people are spending more and more time and money getting from one place to another (Ingram, Carbonell, Hong, & Flint, 2009).

Automobile usage has also been connected with climate change and rising energy costs. Fuel prices have skyrocketed since 2002 and as a result so have household transportation costs. Many people in the U.S. cannot avoid these additional costs, because they do not have adequate alternatives to driving, such as transit.

It is important to note that household use of personal vehicle transit is not the only undesirable aspect of sprawl development. Critics of sprawl also point to the lack of affordable housing and housing options for moderate and low-income households, the destruction of farmland and environmentally sensitive areas, and the negative effects on air and water quality (Ingram et al., 2009). Still, over-reliance on driving is thought to be

a major negative influence of sprawl morphologies. Smart growth advocates have advanced the notion that by changing the built environment to allow for greater density and by creating transportation alternatives and better transit access, people will drive substantially less (Ewing & Cervero, 2001).

Smart growth policies have now been in effect for decades in some jurisdictions. Among the many beneficial effects of smart growth cited by its proponents is the belief that by changing the built environment in which people live, transportation behavior can be altered so that it is more efficient and environmentally friendly. Proponents hope that people living in denser, more walkable neighborhoods will become less dependent on automobiles. Thus, using the available data, it is important to try to discern the influence of the built environment on the way we use transportation in our everyday lives.

By further studying the effects of the built environment on automobile ownership levels and VMTs, the goal of this analysis is to better understand the implications of density and several other spatial variables on household transportation behavior. This is especially important as more and more local governments adopt smart growth planning practices. Since 2001, large Sunbelt cities, such as Houston and Phoenix, which have traditionally seen mostly sprawling automobile-oriented growth, have made substantial public investments in regional rapid transit systems. As these regions begin to promote more compact, transit-oriented neighborhoods, it is important to measure the effect these neighborhoods have on the transportation habits of those who live in them.

A great deal has been written on planning theory and the benefits of smart growth. There are also a number of empirical studies that have looked at the effects of the built

environment on topics related to automobile ownership and VMTs. However, most of these studies rely on very old datasets (Zhang, 2006) or only focus on a single region (Shay & Khattak, 2006). The studies that examined multiple regions do not analyze data beyond the MSA-level scale (Cervero & Murakami, 2010). In studies that have looked at VMTs or automobile ownership at the block group scale, it has typically been within a single MSA or metropolitan area (Haas, Makarewicz, Benedict, & Bernstein, 2008).

This analysis builds on existing research by looking at a sample of 75 block groups across five different MSAs using both neighborhood and regional variables. Because data are examined from multiple regions, the results of this study are more generally applicable than studies that examine a single MSA. Also, the study looks at the influence of variables at overlapping scales (e.g., the neighborhood and the metropolitan region).

Chapter 2: Background

For many years those in favor of smart growth principles have advocated its many benefits. Of particular relevance to this study are the aspects of smart growth that encourage transit use and discourage automobile dependency through changes in the built environment.

Among the most commonly cited examples of smart growth practices for reducing the use of personal vehicles are transit-oriented development, mixed-use zoning, and increased transit options (Gearin, 2004). Transit-oriented development refers to the zoning of high-density development in the immediate vicinity of transit stations. Mixed-use zoning practices allow for the creation of buildings with multiple uses, such as an apartment building with ground-level retail. Mixed-use zoning can also apply at the neighborhood scale where zoning allows for a mix of different building types. The purpose of such mixed-use development practices is to reduce overall transit demand. Many regions increase transit options through the construction of new rail transit systems, the addition of new bus lines, or the creation of bike trails.

Studies have found that the connection between transportation and land-use is complex. Factors, such as household preferences, socio-demographic variables, and the scale at which policies are enacted, determine how the connection manifests itself. Furthermore, while urban form can influence transportation behavior, it is usually a secondary factor to personal preferences and socio-demographics. Increased density, mixed land-use, and transit-oriented design are thought to play only a modest role in decreasing VMTs (Ingram et al., 2009).

This section provides an overview of the current body of literature on the topics relevant to this paper. Many studies have looked into the effect of the built environment on vehicle transit behavior; however, most of these studies only look at a single region and measure a limited number of variables. This study takes a deeper look into the relationship between the built environment and both vehicle miles traveled (VMTs) and automobile ownership by (1) measuring multiple spatial variables, (2) measuring variables at both the block group and metropolitan statistical area (MSA) scale, and (3) collecting data from sample block groups across five MSAs.

2.1 Single Region Studies

A number of studies have looked at the factors that influence transportation behavior in a single metropolitan area. Shay and Khattak (2006) investigated automobile ownership in the Charlotte, NC, metropolitan area, using survey data from 2001. This was done using a number of environmental measures, neighborhood typologies and indices of environmental factors generated by factor and cluster analyses, and other spatial variables. It was shown that automobile ownership is affected by socio-demographic factors such as income and household size. However, environmental factors, including land-use and walkability, had a greater influence on trip generation than socio-demographic factors.

Another approach used in single region studies is to model household transportation expenses for different types of neighborhoods. Haas et al. (2008) developed models for predicting total out-of-pocket household annual transportation expenditures and tested them in the Minneapolis-St. Paul, MN metropolitan area using

data from the 2000-2003 timeframe. These models included independent variables, such as density, job access, neighborhood services, walkability, and transit connectivity. The models were then used to confirm the statistically significant influence of the built environment and transit accessibility on household transportation expenditures. In particular, built environments that featured smaller block sizes, a greater number of services within the neighborhood, greater residential densities, high transit connectivity, and close proximity to major employment centers reduced the number and distance of automobile trips.

The fact that these studies only analyzed a single region severely limits their applicability at larger spatial scales. The likely reason that these studies only tested a single region is that most of them relied on surveys that included data from personal travel diaries. This allowed for the examination of transit behavior at the individual level, but it also limits the applicability of the results to other regions.

2.2 Longitudinal Approaches

Krizek (2003) investigated the effect of neighborhood-scale, urban-form factors on travel behavior in the Central Puget Sound region in the state of Washington. This longitudinal study developed regression models to predict changes in travel behavior as a function of neighborhood accessibility controlling for regional and workplace accessibility. The study used survey data collected between 1989 and 1998. It was found that household travel behavior changed when exposed to differing urban forms and that higher neighborhood accessibility decreases VMTs. Using travel diaries collected in the San Diego, CA, area in 1986, Crane and Crepeau (1998) investigated claims made by

smart growth advocates that urban design can influence travel behavior. The study found that land-use played only a small role in explaining travel behavior. Additionally, the study found no evidence to support the theory that street network patterns affect non-work travel decisions.

2.3 Studies at Multi-Region and Neighborhood Scales

Some studies have investigated VMTs in multiple regions. Cervero and Murakami (2010) analyzed the effects of the built environment on VMTs in 370 urbanized areas in the U.S., using data from 2003. Using structural equation models, factors such as population density, access to employment, population of urbanized area, and rail transit usage, the study found that population density was strongly and positively associated with lower VMTs. Employment access, the population of the urbanized area, and rail usage had only modest effects. This analysis took place at the metropolitan scale as opposed to the block group scale used in this study. This analysis was likely conducted at the MSA scale so that data could be easily compared across multiple regions.

Other studies have investigated transit behavior at the neighborhood scale only. Using 2003 survey data from northern California, Cao, Mokhtarian, and Handy (2007) examined the influence of neighborhood design versus residential self-selection as the causal factor of transportation behavior at the neighborhood level. The study relies on data from four "traditional' neighborhoods and four "suburban" neighborhoods with variables related to travel behavior, neighborhood characteristics, neighborhood preferences, travel attitudes, and socio-demographics. The study found that while

residential self-selection had significant impacts on travel behavior, the built environment also had a statistically significant association with changes in travel behavior. Increased transit accessibility was the most important factor in reducing driving.

Many studies have included only a limited number of spatial measures when investigating vehicle transportation behavior. Zhang (2006) conducted an empirical study of automobile dependence in the Boston area using travel survey data from 1991. The emphasis of the study was mode choice. Spatial variables for land-use and street connectivity were used in addition to socio-economic variables. It was found that automobile dependence is sensitive to street network connectivity and automobile availability. Both population and job density were also found to be important, and land-use's role in increasing transportation options was confirmed.

Cervero (2002) studied the influence of built environments on transportation mode choice using 1994 survey data from Montgomery County, MD. The study developed a normative model that weighed the influence of built environment factors, including density, diversity, and design. It was found that density and mixed land-use had a significant influence on transportation mode choice while urban design factors had a more modest influence.

Kim and Brownstone (2010) looked into the impact of residential density on vehicle usage and fuel consumption. An empirical model was developed using data from the 2001 National Household Travel Survey. It was found that households located in block groups in which density is greater than a 1,000 housing units per square mile will drive less and consume fewer gallons of fuel than households in less dense areas.

Density in the context of the surrounding area was found to be greater than the effect of just residential density. It was also found that moving a household from a suburban area to an urban area reduces household VMTs by 15%. While the Kim and Brownstone (2010) study uses the same VMT data used in this analysis, it only examined VMTs in the context of density.

Hess and Ong (2002) used 1994 survey data from Portland, Oregon, to develop a model to explain automobile ownership based on demographic variables and a few urban design characteristics such as land-use mix. It was found that the presence of mixed land-uses caused the probability of owning an automobile to decrease by 31 percent. It was also found that non-sprawl neighborhoods are more conducive to walking and to the use of public transit.

Some studies have used methodologies similar to the one used in this analysis. Cervero (1996) investigated the suggestion that mixed land uses encourage non-auto commuting. His analysis used data from the 1985 American Housing Survey that included survey data for eleven MSAs. A regression model was created using various land-use variables designed to capture the density and presence of mixed land-use. Control variables measured household income, automobile ownership, location within MSA, the presence and adequacy of transportation choices in neighborhood, and the distance from home to work. The study found that neighborhood density had a greater influence than mixed land-uses in influencing transportation choices, except for walking and bicycling. The study also found significant elasticity between land-use environments and commuting choices in the eleven MSAs.

Using variables that describe the built environment, this study seeks to explain their effect on vehicle transit behavior. By measuring multiple spatial variables at both the block group and MSA scale across five MSAs, this study provides a deeper analysis of both VMTs and automobile ownership than previous studies.

Chapter 3: Methodology

This study seeks to examine how built environments at neighborhood and regional scales relate to transportation behavior. The analysis focuses on two variables that measure transportation behavior at the block group level: automobile ownership and vehicle miles traveled (VMTs). The influence of the built environment on these variables is examined across a range of cases in metropolitan statistical areas throughout the United States, representing both neighborhoods with a dense residential population and in less dense "sprawl" neighborhoods.

The research examines vehicle transit behaviors using a cross-sectional approach for block groups from five regions for the 2000-2001 timeframe. The study measures and examines several independent variables related to the built environment at the scale of neighborhoods and metropolitan regions, including density, job access, transit access, retail access, and street connectivity. Through analysis of all of these variables, this paper attempts to find the relationships that explain household automobile ownership levels and VMTs.

This chapter serves as an overview of the steps taken to complete this analysis, including the sampling framework, major hypotheses, variables, and spatial and statistical modeling processes. All census data used in this study comes from the 2000 U.S. Census.

3.1 Sampling Framework

This study samples seventy-five block groups, combining stratified random and selected sampling methods. The block group is the smallest scale at which detailed

information is available for most of the key variables in this study. This unit of analysis is small enough that in suburban and urban areas it covers a geographic area similar to what might be termed a neighborhood. However, in less-populated, exurban areas the geographic area of a block group is much larger.

The seventy-five cases are sampled from five metropolitan areas in different regions of the United States. These metropolitan areas are Chicago, IL; Miami, FL; Portland, OR; San Diego, CA; and Washington, DC (Figure 1). These areas were selected because they each represent a different region of the U.S. They are also diverse in terms of their overall population, urban form, the era in which they developed, and the extent of their regional transit network.

Figure 1: Sample Metropolitan Statistical Areas (MSAs)

Throughout this study, metropolitan areas are defined by their metropolitan statistical area (MSA). MSAs are defined by the U.S. Office of Management and Budget. They are composed of counties or county-equivalents that cover the extent of a central urban area or urban cluster and the surrounding area of somewhat continuous and relatively high population density. Additional "outlying counties" that have strong economic ties to the central MSA are also included in an MSA. These outlying counties are included in an MSA if the total in-commuting and out-commuting (i.e., the employment interchange) exceeds 25% of the total employment in the outlying county.

The MSA is used as the regional unit of analysis in this study because it is the only consistent definition of a metropolitan area that is available. It is also the definition of metropolitan areas used by the census when discussing large cities and their surrounding suburbs (United States, 2000).

The Chicago MSA (Figure 2) was the third most populated region in the 2000 U.S. Census and has an extensive transit network. It is the largest in the sample group of MSAs and developed earlier than all but Washington, DC. While the built environment in the city of Chicago is characterized by traditional, dense urban development, most of its surrounding suburbs are sprawling and automobile-oriented. However, some of the older inner-suburbs are more dense and transit-oriented.

The Miami MSA (Figure 3) was the sixth largest in the 2000 U.S. Census and has a relatively small transit infrastructure. While development in the coastal areas is highly dense, the vast majority of the region is characterized by low-density automobile-dependent development typical of most Sunbelt cities.

Ranking twenty-fifth in the 2000 U.S. Census, the Portland MSA (Figure 4) is by far the least populated MSA included in this study. However, the Portland area has a relatively extensive transit system for its size. Portland was a pioneer in adopting smart growth principles; as a result, it is denser than MSAs of comparable size.

Figure 2: Chicago, IL MSA
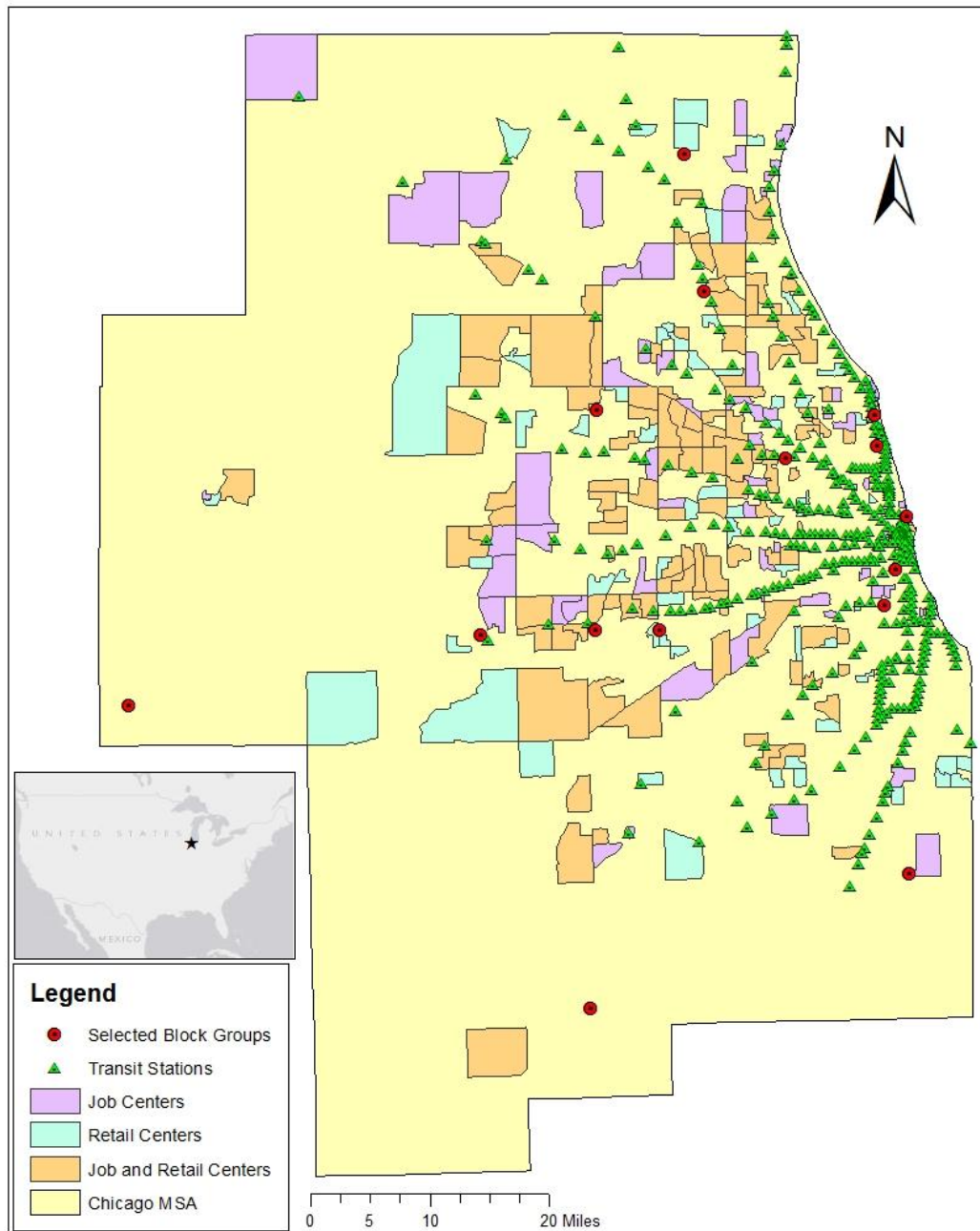
The San Diego MSA (Figure 5) ranked seventeenth in population in the 2000 U.S. Census. It has a transit system of moderate extent. The built environment in most of the region is less dense than other MSAs in this study.

Figure 3: Miami, FL MSA
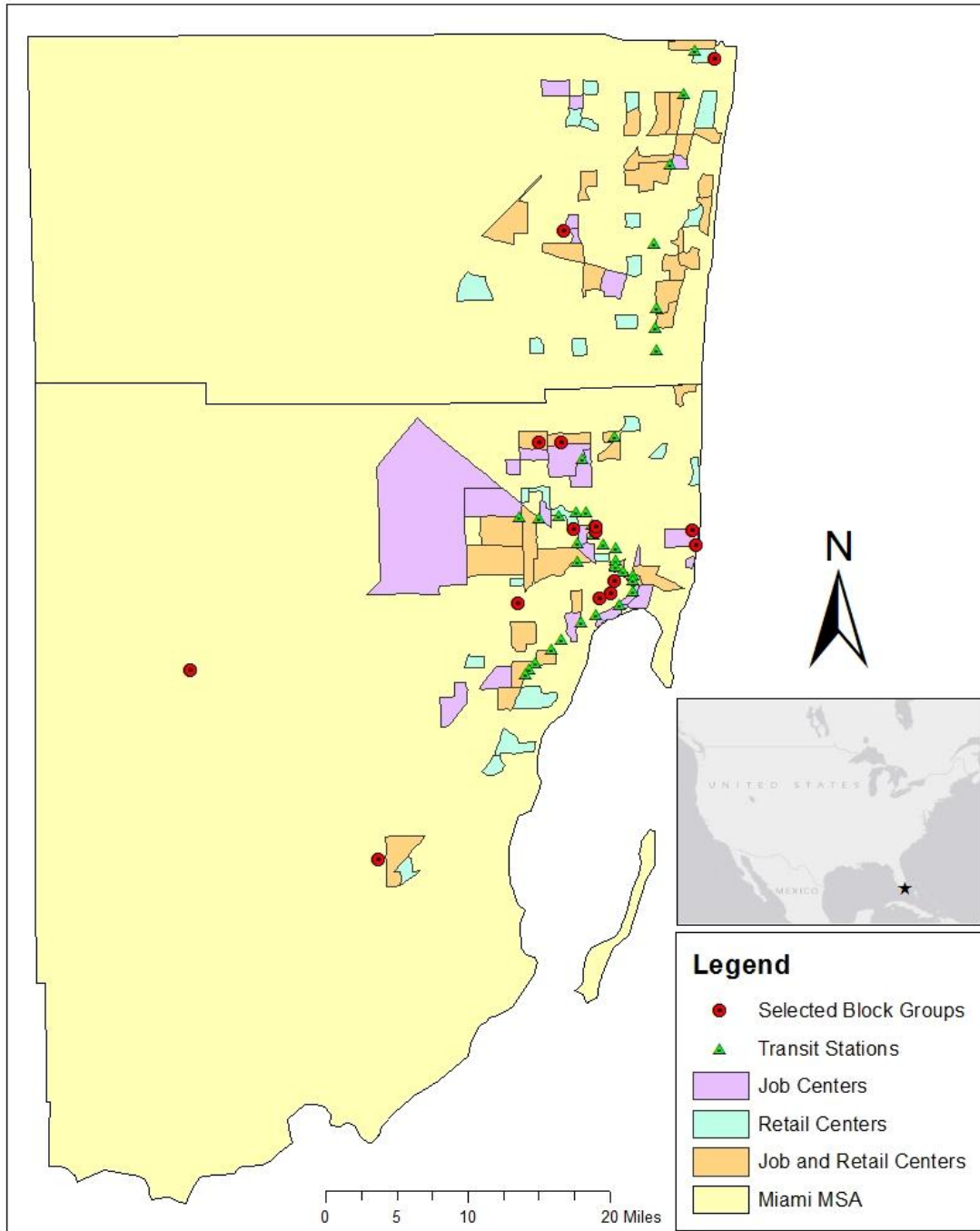
Figure 4: Portland, OR MSA
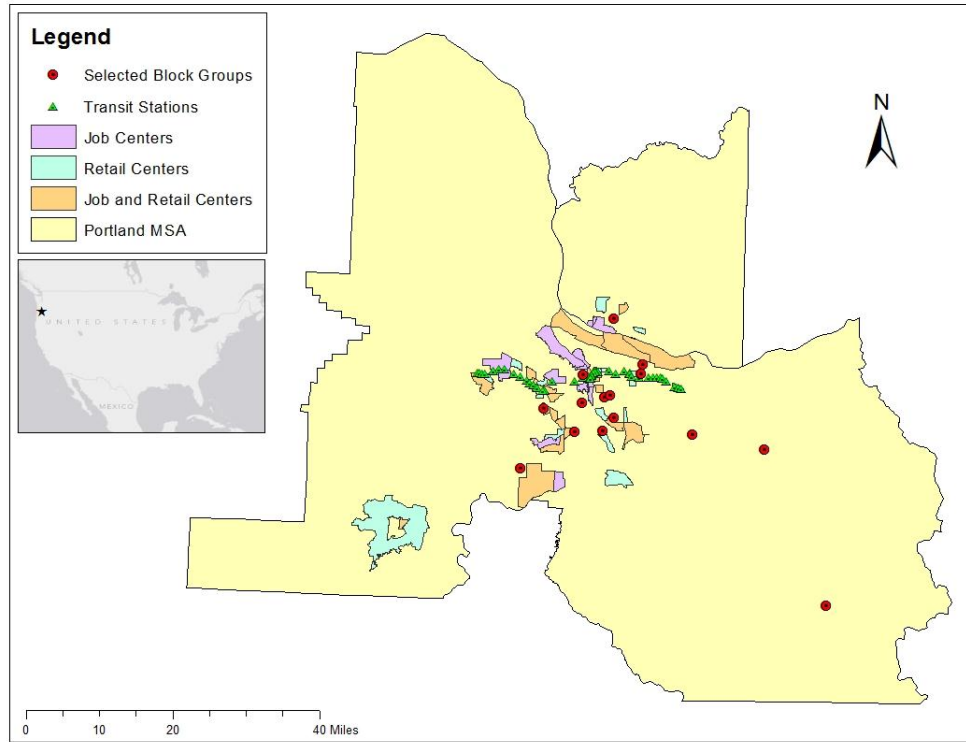
Figure 5: San Diego, CA MSA
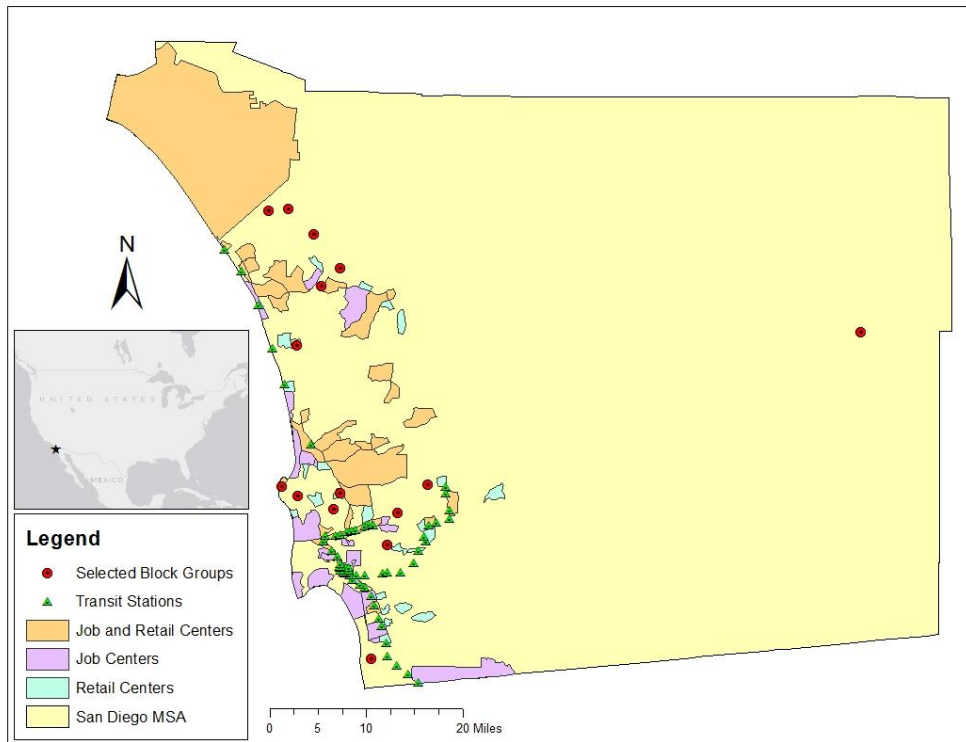
The Washington, DC MSA (Figure 6) was the seventh most populated MSA in the 2000 U.S. Census. It has a relatively large transit infrastructure, and jurisdictions in the Maryland suburbs have been subject to statewide smart growth initiatives since 1997 (Ingram et al., 2009). As a result, many suburban areas have areas of high density in the vicinity of transit stations.

Figure 6: Washington, DC MSA

As shown in each of the maps above, fifteen sample block groups are taken from each metropolitan area of which thirteen are randomly selected. Additionally, the most dense and least dense block groups within the MSA are included in the sample.

3.2 Major Hypotheses and Dependent Variables

It is expected that the data will show that the built environment influences the number of automobiles owned per household (Table 1). Furthermore, it is predicted that automobile ownership will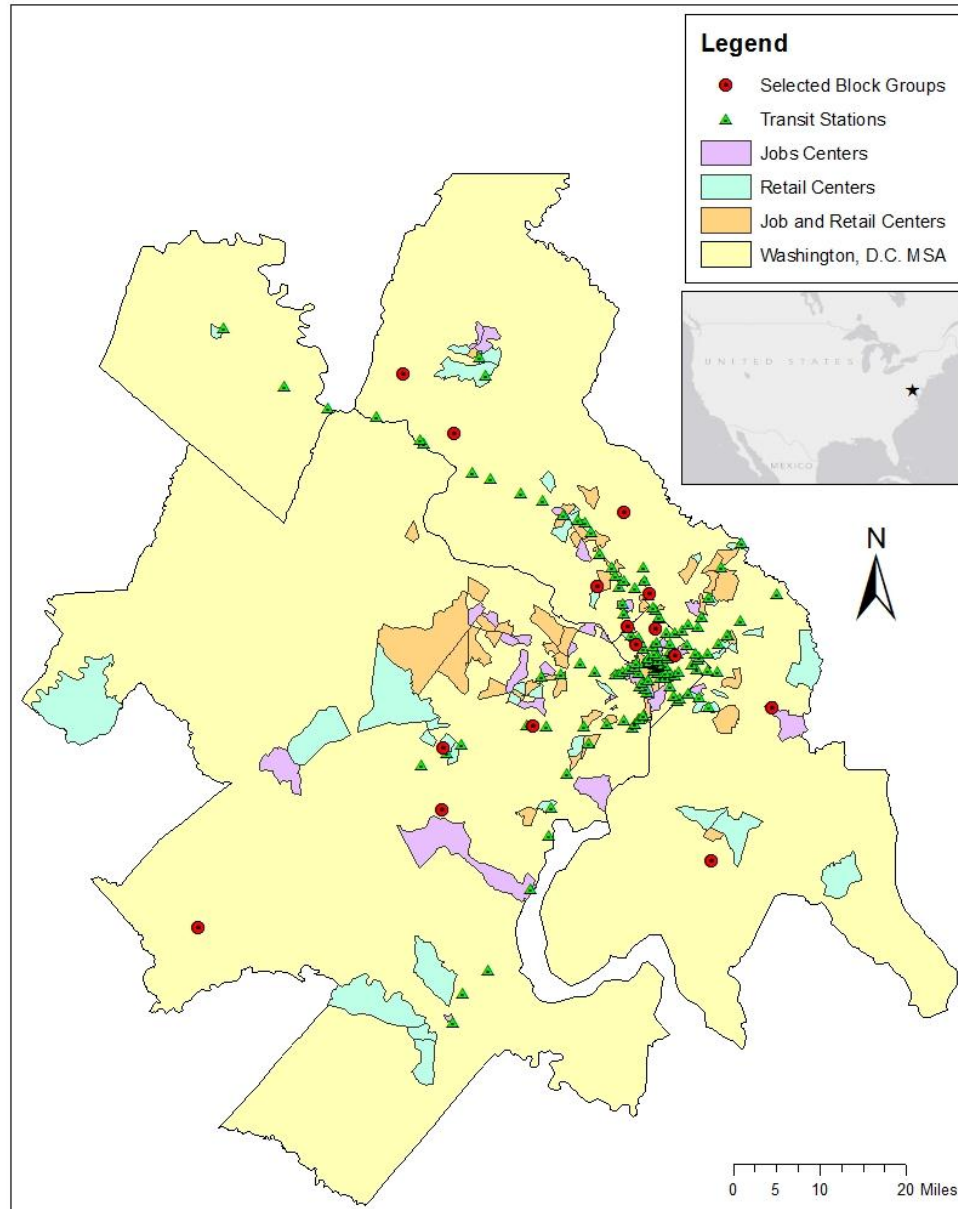 be greater in sprawl neighborhoods and lesser in denser neighborhoods. The data used to measure automobile ownership comes from Summary File 3 of the 2000 U.S. Census (United States Census, 2000). The census variable used describes the total number of vehicles available in a block group. To measure automobile ownership by household, this value is then divided by the number of households in that block group. The result is a ratio measure that describes the average number of automobiles owned per household within a block group. Similar studies have used census data to measure automobile ownership (Center for Transit-Oriented Development, 2006). The number of automobiles owned by a household can be indicative of the automobile-dependence of the individuals who live within it. Less automobile-dependent households may only require a single car among its occupants, while the occupants of other households may each require a car for their daily transportation needs.

Table 1: Dependent Variables

Variable	Source	Hypothesis
Automobile Ownership	2000 U.S. Census	The built environment influences the number of automobiles owned per household
Vehicle Miles Traveled	2001 National Household Travel Survey	The built environment influences household VMTs

It is also predicted that the built environment influences driving such that VMTs are greater in sprawl neighborhoods and lesser in more densely populated neighborhoods. The data used to measure VMTs came from the 2001 National Household Travel Survey (NHTS) that was conducted by the Federal Highway Administration. NHTS data is reported annually at the MSA, state, and national scale. VMTs are a strong indicator of how far individuals are traveling daily to access jobs and amenities.

While data from the NHTS is typically unavailable at any scale below the MSA level, a model was developed to estimate the data at the census tract level using data from the 2001 survey. The variables used to model VMTs by census tract were household size, household income, and employment rate (Hu, Reuscher, Schmoyer, & Chin, 2007).

Unfortunately, block group scale data for VMTs do not exist nationally and would require significant time and expense to collect. The census tract level is the smallest scale at which VMT data have been estimated across multiple MSAs. Furthermore, the NHTS estimation model cannot be used to downscale the survey data to the block group level, because the employment rate variable used in the model is not available at scales smaller than the tract level.

Therefore, sample block groups used in this study are assigned the VMT value of their corresponding census tract. All VMT estimates are per household on an average weekday. The tract level estimates for VMTs are given based on household size and the number of vehicles available to a household. For the purposes of this study, the VMT estimate for each block group is based on its mean household size and the mean number of automobiles per household.

3.3 Measuring the Built Environment: Independent Variables

Dense neighborhoods are almost always less automobile friendly than less dense areas, in part because the availability of parking is scarce. Therefore, density will likely affect both VMTs and automobile ownership. Density of the built environment is assessed as the number of households in a block group divided by the total acreage of the block group. Measuring households per acre as opposed to population per acre is a more appropriate method for this study because it is more closely related to the density of the built environment rather than the population density of a neighborhood. That is, high residential density likely indicates the presence of multi-family housing structures. These buildings contain multiple units that are smaller than a typical single-family home and, therefore, are likely to have fewer occupants per household. If measured by population rather than by the number of households, this could lead to an underestimation of density in such neighborhoods. A negative relationship is expected between density and both VMTs and automobile ownership levels (Table 2).

Table 2: Independent Variables

Variable	Source	Hypothesis
Density	2000 U.S. Census	Expected to have negative relationship with VMTs and automobile ownership.
Job Access	2000 Census Transportation Planning Package	Expected to have negative relationship with VMTs and automobile ownership.
Retail Access	2000 Census Transportation Planning Package	Expected to have negative relationship with VMTs and automobile ownership.
Transit Access	2011 National Transportation Atlas Database	Expected to have negative relationship with VMTs and automobile ownership.
Street Connectivity	2000 U.S. Census	Expected to have negative relationship with VMTs and automobile ownership.

The commute to and from work is a key aspect of people's daily transportation behavior. The distance between a person and their job is likely a major factor in the decision of how to get to work   Job access is measured as the distance from a block group to an "employment center." For the purposes of this study, an employment center is defined as a census tract that is among the top ten percent within an MSA in terms of the total number of jobs. Similar measurements have been used by authors to establish the location of job centers within a region in studies that attempt to estimate household transportation costs (Haas et al., 2008).   All job data are extracted from the 2000 Census Transportation Planning Package (CTPP). A block group's distance from the nearest employment center could suggest how far people are traveling to get to work. The distance between sample block groups and employment centers is measured from the centroids of both polygons. A negative relationship is expected between job access and both household and automobile ownership levels and vehicle miles traveled.

Similar to a workplace, retail services, such as grocery stores and shopping malls, are accessed by many people on a frequent basis. The distance to these services likely affects the mode of transportation used to access them. To measure retail access, the distance between a sample block group and a retail center is measured. Those census tracts that are among the top ten percent in an MSA in terms of the aggregate number of retail jobs are considered "retail centers." Retail job data comes from the 2000 CTPP. The measurement is made from the centroid of the sample block group to the centroid of the retail center tract. A negative relationship is expected with both VMTs and automobile ownership per household.

Transit is only an option for individuals if it is reasonably accessible. If it is not easily accessible, it is less likely that a person will choose transit as their primary means of transportation. Transit access is a measurement of the distance between a sample block group's centroid and the nearest transit station. This measure predicts the ease of access to a transit system in a given census block group. Households located near transit stations likely own fewer cars and travel shorter distances. A negative relationship is expected with both VMTs and automobile ownership per household.

This study includes transit stations that are part of a regional light, heavy, or commuter rail system. Unfortunately, no national dataset could be found that included bus stop data for local or regional bus services. The spatial data for rail transit stations comes from the 2011 National Transportation Atlas Database that uses data from the U.S. Department of Transportation. In keeping with the timeframe of this study, only stations that were in operation in 2000 are included.

Street connectivity is one relatively simple measure of the ease of walking, as it is related to pedestrian connectivity in a neighborhood. The street connectivity of a neighborhood is thought to influence transportation mode choice because people are willing to walk greater distances more frequently in areas with high pedestrian connectivity (Center for Neighborhood Technology, 2010).

Street connectivity is measured as the block group's acreage divided by the number of blocks within the block group. This measurement gives an indication of how complete the street network is in a sample block group. A street network can be considered as a proxy for a sidewalk network. Shorter block lengths have been shown to

encourage pedestrian activity (Huang, Stinchcomb, Pickle, Dill, & Berrigan, 2009). A negative relationship is expected between street connectivity and both automobile ownership and VMTs.

3.4 Confounding Variables

This study also tests against a series of confounding variables to explore whether variables, not from the built environment or at spatial scales larger than the block group, are related to block groups and vary in such a way as to confuse or suppress the observed relationships. The confounding variables tested are household income, regional density, extent of regional transit network, age of neighborhood population, and individual transit expenditure (Table 3).

Income is measured as the per capita household income of a block group. Household income has been shown to influence transportation mode choices. Higher income households are more likely to own automobiles and use transit less. This measure has been used in other models to explain transportation choices (Haas et al., 2008). A positive relationship is expected between income and both automobile ownership and VMTs.

Regional density is measured at the MSA level as the MSA population per square mile. Dense metropolitan areas may lead to greater use of transit options than in less dense regions. Because this variable is more oriented toward a measurement of population density rather than strictly the density of the built environment, population per square mile is more appropriate than using households. Regional density is expected to relate negatively to both VMTs and automobile ownership.

Table 3: Confounding Variables

Variable	Source	Hypothesis
Household Income	2000 U.S. Census	Positive relationship expected with both automobile ownership and VMTs.
Regional Density	2000 U.S. Census	Negative relationship expected with both automobile ownership and VMTs.
Extent of Regional Transit Network	2000 U.S. Census, 2011 National Transportation Atlas Database	Negative relationship expected with both automobile ownership and VMTs.
Age of Neighborhood Population	2000 U.S. Census	Negative relationship expected with both automobile ownership and VMTs.
Individual Transit Expenditure	2000-2001 Consumer Expenditure Survey	Negative relationship expected with both automobile ownership and VMTs.

This study also measures the extent of a region's transit network. Overall, people are likely less dependent on automobiles in regions in which there is a more robust transit network. This variable is measured at the MSA level as the number of residents per heavy, light, or commuter rail station. This variable is obtained by dividing the MSA population by the number of transit stations in the MSA. The extent of a transit network in a region is expected to relate negatively to VMTs and automobile ownership.

The age of neighborhood population variable is measured at the block group level as the median age of residents. An individual's age may affect the transportation options that are available to them. The median age of a neighborhood is expected to relate significantly in the case of outliers of young or old neighborhoods. In areas with more children or a high elderly population, people likely own fewer automobiles.

Individual transit expenditure is measured at the MSA level as the average annual consumer expenditure on public transportation. This variable could provide an indication of how people are using regional public transportation systems based on the average annual dollar amount spent on transit. Data for this variable come from the Bureau of

Labor Statistics' 2000-2001 Consumer Expenditure Survey. A negative relationship is expected between individual transit expenditure in both automobile ownership and VMTs.

3.5 Spatial Modeling of Variables

All spatial data were processed and analyzed using ArcGIS 10. The analysis uses U.S. Census TIGER/Line shapefiles representing MSAs, block groups, tracts, and blocks as defined in the 2000 U.S. Census. Census demographic data, in addition to NHTS and CTPP data, were merged with their corresponding shapefiles at the tract and block group level within ArcMap. The merge was based on the census-designated tract or block group identification number. Throughout this study, distance was always a measurement of distance via the road network as opposed to "as the crow flies." Additionally, distances were always measured from the centroid of a block group or tract polygon.

The spatial modeling process began by creating the sample of fifteen block groups for each MSA. The U.S. Census American Fact Finder website was used to create an Excel spreadsheet of all block groups within a selected MSA. Once obtained, a "Random" field was added to the spreadsheet, and a random number generator was used to assign random values between 0 and 1. The spreadsheet was then sorted based on the "Random" field, and the thirteen block groups with the lowest random number were then used as the randomly sampled block groups from their MSA. The last two block groups identified for the sample were the most and least dense block groups within the MSA, to reach a total of 15 for each MSA.

To build the sampling frame, a shapefile with all U.S. block groups was added in ArcMap along with a shapefile of all MSAs. A clip operation of the block groups layer using the MSA layer was performed to create a shapefile of just the block groups that fall within the MSA.

The clipped block groups layer was then projected using the appropriate state plane projection for the region. Using the calculate geometry tool, the acreage of each block group was calculated and added to the table. A table containing general block group level demographic data was then merged with the block group data. Among these new data was a variable containing the total number of households within each block group. The field calculator tool in ArcMap was used to calculate the variable for density by dividing the total number of households in a block group by its acreage. This new attribute was the measure of density used in this study. The block group with the highest value was added to the sample as the most dense block group within the MSA. The block group with the lowest value that had at least 100 households was added to the sample as the least dense block group within the MSA.

For the dependent variables, automobile ownership per household was obtained by merging the census data table containing the total automobiles owned in each block group. A field was then added to the table for automobile ownership per block group, and the field calculator tool was used to divide total automobiles by total households for each block group.

For block group VMTs, NHTS estimates for the corresponding census tracts were used. The NHTS estimates for each tract varied depending on the size of a household

and the number of vehicles owned in that household. For the purposes of this study, the mean household size and automobile ownership value of a sample block group was used to determine the NHTS estimate of VMTs in its corresponding census tract. The census tract level estimate of VMTs was then applied to its corresponding block group.

The process of determining job and retail access began by identifying the job and retail centers within each MSA. The CTPP data for the locations of jobs were presented at the tract level. The data were sorted by the field representing total jobs within each tract. All tracts except those within the top 10% of total jobs were deleted. The table was then brought into ArcMap and merged with the tract shapefile for the MSA, creating a job centers shapefile. The same process was used for identifying retail centers. Retail and job access were measured as the distance between the centroid of a sample block group and the centroid of a retail and job center, respectively. All centroids were created in ArcMap using the feature-to-point tool.

The measure of distance for retail access, job access, and transit access was calculated on the road network. To do this, a road network was created within ArcMap. ESRI's detailed roads shapefile available for download through ArcGIS Online was used to create the street network. Once downloaded, the street layer was clipped using the MSA shapefile in order to make the file size more manageable. In ArcCatalog, a new street network was created using the detailed roads shapefile. The new network took into account roadway elevations (i.e., grade-separated intersections) and used length as a constraint. The default options were used for all other settings.

Once the street network was added to ArcMap, "a new route" was created. Using the network location tool, a location was created on top of a centroid. A second location was placed on the centroid of the nearest job center, retail center, or transit station. A new route was then created on the street network between the two network locations. The distance was then recorded. If there were several candidates for the closest job center, retail center, or transit station, multiple measurements were taken to determine the closest. This process was repeated for all selected block groups in order to determine its closest job center, retail center, and transit station.

Street connectivity was measured as a block group's total acreage divided by the number of census blocks in the block group. To measure street connectivity, a shapefile containing the blocks within an MSA was added into ArcMap. An SQL query was created using the Select by Attribute tool to select all blocks that share the same block group identification as the selected block groups. The total number of blocks in each of the selected block groups was counted and added to the selected block groups attribute table. A new street connectivity field was then created in the selected block groups table. Data were populated by using the field calculator to divide the total acreage of the block group by the total number of blocks.

For the confounding variables, the data for block group per capita household income were taken directly from the census. To measure regional density, the MSA shapefile was projected into the appropriate state plane coordinate system. The square mileage of the MSA was then measured using the calculate geography tool. Total MSA population data were taken from the census and added to the selected block groups data

table. The Field Calculator was used to calculate the MSA total population divided by the MSA square mileage.

3.6 Statistical Analysis

Using both the SPSS and R statistics packages for processing, the best fitting multivariate regression models were developed for each dependent variable. To accomplish this, descriptive statistics, such as mean, standard deviation, and histograms were drawn for each variable. Skewness of variables was observed. A full correlation matrix of variables was also drawn to check for multi-colinearity. As detailed in the Results section, the data were further explored for both spatial autocorrelation and the modifiable areal unit problem (MAUP). Various hypotheses were confirmed or rejected as related in the Results and Discussion and Conclusion sections.

Chapter 4: Results

This chapter examines the results of the analysis. It covers the following topics: (1) analysis of study variables, (2) assessment of correlation in study variables, (3) regression models, (4) spatial autocorrelation, and (5) the modifiable areal unit problem.

4.1 Analysis of Study Variables

Households from the sample block groups have a mean VMT value of 50 miles on an average weekday (Figure 7) with a standard deviation of 22. The means for all of the MSAs fall within the standard deviation. The most dense block groups have a mean VMT value of 21 miles while the least dense block groups have a mean of 73 miles.

Figure 7: Average Household VMTs
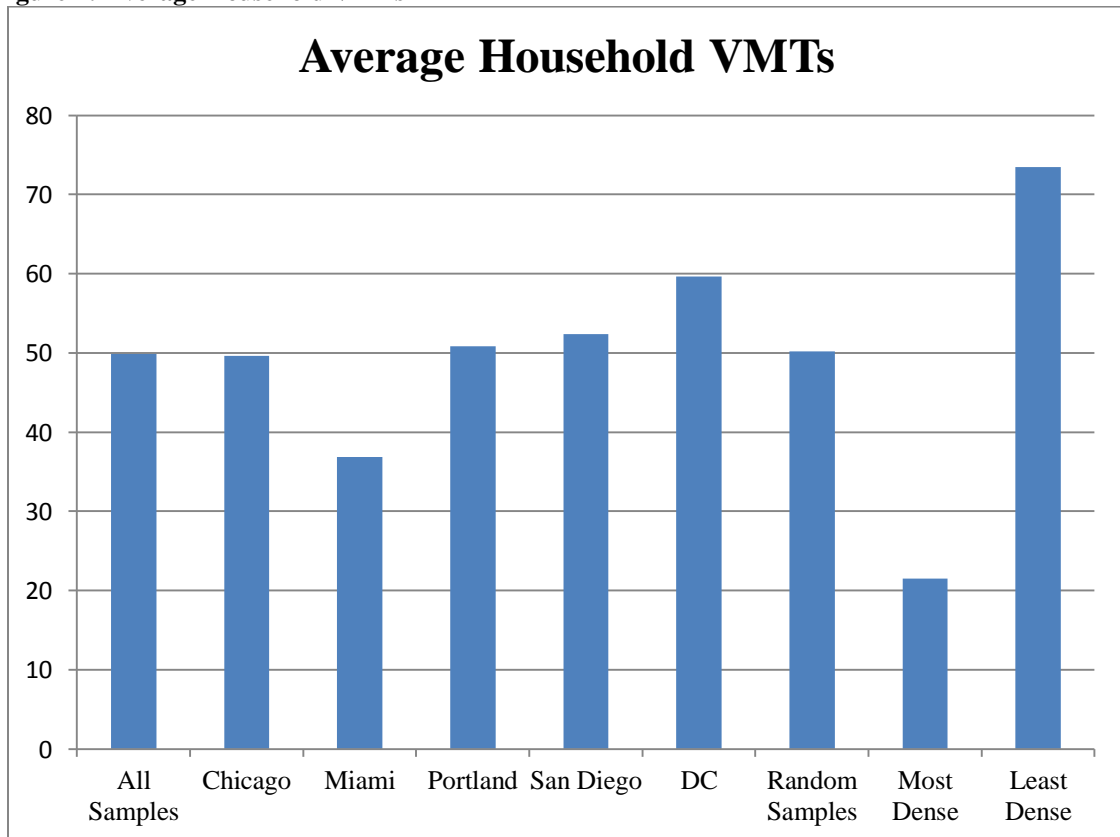
The mean number of automobiles owned per household in the sample block groups is 1.68 (Figure 8) with a standard deviation of 0.5. The mean automobile ownership does not vary greatly in any of the MSAs. The most dense block groups own a mean of .79 automobiles while the least dense own a mean of 2.23. For both dependent variables, the means are similar across all five MSAs, and the most and least dense block groups vary as expected.

Both dependent variables are normally distributed (Table 4). On average, the sample block groups are located 5.72 miles from job centers, 5.66 miles from retail centers, and 7.43 miles from the nearest transit station. The values for street connectivity range from 1 to 2,920 acres per block with the lowest values signifying block groups estimated to have the greatest pedestrian connectivity. The average street connectivity value is 129 acres per block. Natural log transformations are used for all of the independent variables when they are included in a model. This is a common practice in regression modeling to standardize skewed variables and improve model fit (Allison, 1999).

In the independent variable histograms, the most and least dense block groups are significant outliers that cause the histograms to skew right. For the job access and retail access variables, over 85% of the block groups are less than ten miles from a job center and retail center. The remaining block groups (including all five least dense block groups) are between 10 to 61.3 miles from the closest job or retail center. The variable for transit access is similar to the job and retail access variables in terms of distribution with over 82% of the block groups located within ten miles of a transit station. The same

block groups that were outliers for the job and retail access variables are among the 13 that are more than 10 miles from a transit station. The highest value for transit access is 65 miles.

Both the density and street connectivity variables are highly clustered. Only seven of the 75 sample block groups had a density value of more than twenty households per acre. Of the seven, five were the most dense block groups selected from each MSA. The other two block groups are in the San Diego MSA. For street connectivity only 6 block groups had a value of more than 250 acres per block. Of these six, five are the least dense block groups in each MSA. The sixth block group is in the Washington, DC MSA.

Figure 8: Average Automobile Ownership
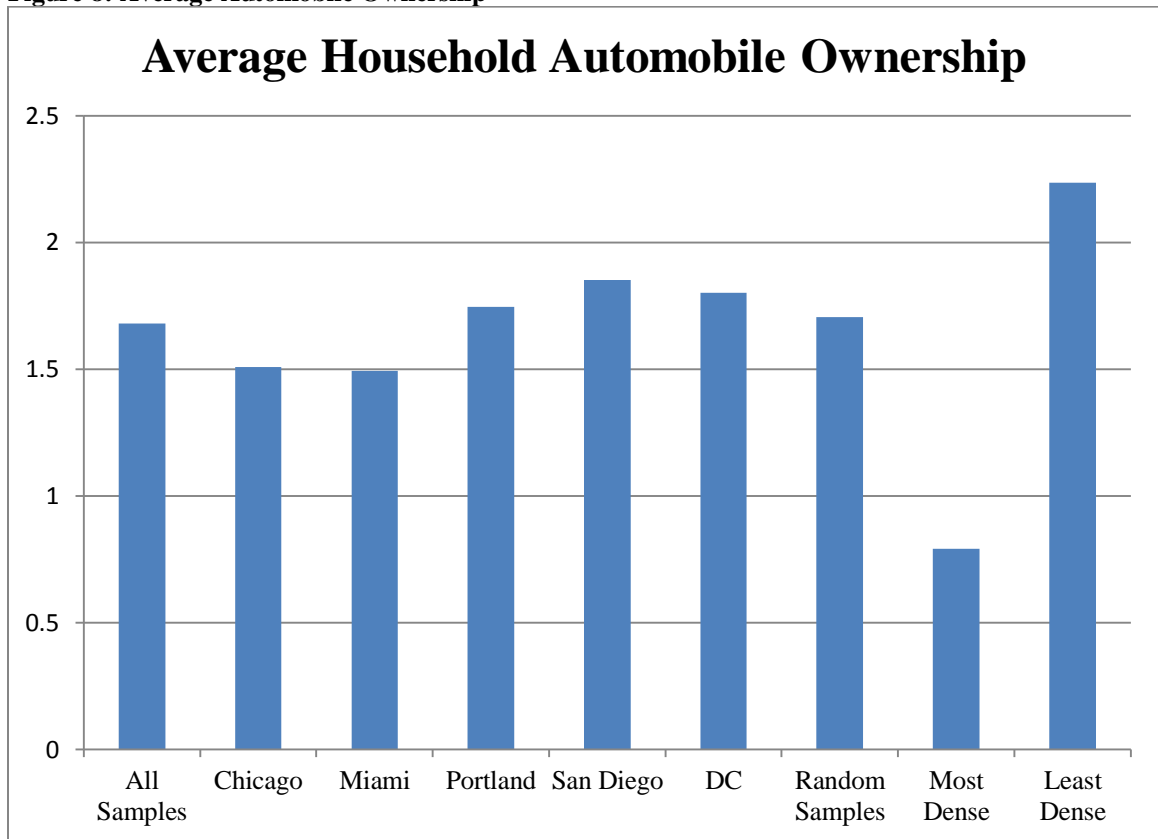
The average block group density is 13.36 households per acre. The average per capita household income is $24,390. The average of the median age of the block group's population is 36.52.

The variables for individual transit expenditure, regional density, and extent of a regional transit network are all MSA-level variables. Therefore, there are only five unique values for these variables. The MSA averages for annual transit expenditure range from $385-$861 per person. The MSA averages for regional density range from 373.67-1,610.66. The average number of citizens per transit station ranges from 21,212-117,466.

Table 4: Distribution of Variables

Variable	Distribution	Natural Log Transformation Used in Models
Dependent		
Automobile Ownership	Normal	No
Vehicle Miles Traveled	Normal	No
Independent		
Density	Skewed Right	Yes
Job Access	Skewed Right	Yes
Retail Access	Skewed Right	Yes
Transit Access	Skewed Right	Yes
Street Connectivity	Skewed Right	Yes
Confounding		
Household Income	Normal	No
Regional Density	Normal	No
Extent of Regional Transit Network	Normal	No
Individual Transit Expenditure	Normal	No
Age of Neighborhood Population	Normal	No

4.2 Assessing Correlation of Study Variables

A significant challenge in building the best-fit models for VMTs and automobile ownership is that many of the independent variables are significantly correlated with one another. In particular, distance to job center and distance to retail center were found to correlate at Pearson's R=.797 (p<0.01 level of significance). The strong correlation between job access and retail access is likely due to the difficulty of distinguishing between the two variables spatially. Tracts that were designated as a job center or a retail center were often both a job and a retail center. In every MSA, combined job and retail center tracts were more common than separate tracts for job centers or retail centers (Table 5).

Table 5: Breakdown of Spatial Relationship between Job and Retail Centers

MSA	Job Centers	Retail Centers	Job and Retail Centers
Chicago	68	68	120
Miami	26	26	36
Portland	17	17	26
San Diego	25	25	36
D.C.	47	47	57
Total	183	183	275

Job access and retail access were not significant in either the VMT or automobile ownership regression models. An additional variable was created for the mean of the distance to job center and retail center variables for each block group. This variable was also not significant in either model. Although job access and retail access were also highly correlated with transit access, the variable for distance to transit was consistently more significant in test models and therefore was selected over the other distance

measures. Because the three distance variables were highly correlated, including more than one in a regression model did not increase the model's explanatory power.

All independent variables were found to have significant bi-variate correlations with both dependent variables (p≤0.01). Of the confounding variables, transit spending and transit stations per capita showed a bi-variate correlation with VMTs at the p≤0.05 level of significance. For automobile ownership, only MSA density was correlated at the p≤0.01 level of significance.

4.3 Regression Models

A best-fit linear regression model was developed for both dependent variables. The VMT model includes two variables—the natural log transformation of block group density and the extent of a region's transit network (Table 6). Adjusted R-squared is used in this study to compare the predictive power of different models. For models that have a low number of samples and many predictor variables, adjusted R-squared minimizes bias (Agresti, 2009). The adjusted R-squared for the model is .404 meaning that about 40% of the variation is explained by density and the extent of the regional transit network. The model shows that as density increased, VMTs decreased. Additionally, as the number of residents per transit station increased, VMTs also increased.

The transit access variable was not included in this model but was very close to having a statistically significant association with VMTs. The job access, retail access, and street connectivity variables showed no significant association with VMTs when included in models. Household income could not be included in the model because

income was a variable used to estimate tract-level VMTs from the 2001 National
Highway Travel Survey (NHTS) data.

A model was also developed for VMTs using only the randomly selected block
groups, thereby removing induced bias from the sample. In addition to the block group
density and the extent of a region's transit network, distance to the nearest transit station
was found to be statistically significant and therefore is included in this model. The
adjusted R-squared of the second model was .459.

Table 6: Best-Fit Regression Model for VMTs

Vehicle Miles Traveled		
Adjusted R-Squared	**0.404**	
Model	**Intercept**	**Significance**
Constant	17.396	0
Block Group Density	-6.524	0
Extent of Regions Transit Network	-2.85	0.006
N=75		

In the best-fit model for automobile ownership, three variables were found to be
significant: (1) block group density, (2) distance to nearest transit station, and (3) per
capita household income. The adjusted R-squared of this model was .445 (Table 7). The
results of the model showed that higher block group density was related to a decrease in
automobile ownership. As the distance to the nearest transit station increased, so did
automobile ownership. Similarly, as household income increased, automobile ownership
increased.

The removal of induced bias in the sample does not significantly change the
model. A second model was developed using only the random samples. This model used
the same variables as the first automobile ownership model, and the direction of each

variable was the same as in the previous model. The second model's adjusted R-squared was .479.

As with the VMT model, access to job centers and retail centers has no significant association with automobile ownership. Street connectivity also shows no significance in the automobile ownership model. This would seem to confirm an association between the smart growth principles of transit access and density with automobile use. Not surprisingly, the model results also indicate an association between wealth and automobile ownership. This indicates that wealthier households are more likely to own more than one car and presumably drive more even when controlling for built environment factors.

The best-fit regression model for vehicle miles traveled (VMTs) shows that the variables for density and the extent of a region's transit network explain over 40% of average weekday VMTs per household. When the induced bias introduced by including the most and least dense block groups in each MSA is removed, the model explains nearly 46% of average weekday VMTs per household. When working with just the random samples, the distance of a block group to the nearest transit station becomes statistically significant and is therefore included in the model.

Table 7: Best-Fit Regression Model for Automobile Ownership

Automobile Ownership		
Adjusted R-Squared	**0.445**	
Model	**Intercept**	**Significance**
Constant	12.061	0
Block Group Density	-4.437	0
Distance to Transit	1.999	0.049
Household Income	2.105	0.039
N=75		

Additionally, the hypothesis that the built environment influences the number of automobiles owned per household is supported by the data results. For automobile ownership, the best-fit regression model indicates that density, distance to transit, and household income explain over 44% of household automobile ownership. With the removal of the most and least dense block groups that create induced bias, the same variables explain nearly 48% of household automobile ownership rates.

Several of the independent and confounding variables do not increase the explanatory power of either model and, therefore, are not included. As mentioned above, the variables for job and retail access are not included because of their high correlation with transit access. Street connectivity is correlated with both VMTs and automobile ownership at the $p \leq 0.05$ level of significance. However, it is not significant when included in either model.

For the confounding variables, the variables for regional density, age of neighborhood population, and individual transit expenditure are omitted from both models. Regional density does not have a bi-variate correlation with VMTs, but it does with automobile ownership at the $p \leq 0.05$ level of significance. The age of neighborhood population variable does not have a bi-variate correlation with either VMTs or automobile ownership. Individual transit expenditure has a bi-variate correlation with VMTs at the $p \leq 0.05$ level of significance. It does not correlate with automobile ownership. When included in either model, these variables were not found to be significant.

4.4 Spatial Autocorrelation

Spatial autocorrelation measures the relationship of a variable among multiple occurrences in space (O'Sullivan & Unwin, 2010). In some cases, data from locations within close proximity may be more likely to be similar than data from further locations. Thus, adjacent location, rather than the study variables, may explain or enhance statistical associations between block groups and the dependent variables, either within or between regions.

To test for autocorrelation of the sample block groups within each MSA, two approaches have been used. First, contrast models were developed for all the MSAs to test for the potential significance of variables not yet measured at the regional level. Second, comparisons of fixed effect and mixed effect models were developed to test the potential of the MSAs and the regional level variables to explain residual variance. Third, the X and Y coordinates of the centroids for each of the sample block groups were measured and entered into the regression models to test for association of proximity of the block groups with the dependent variables within each region.

For the contrast model, a dummy variable has been created for each MSA to test for the significance of a particular MSA compared to all the others (i.e., the test region=1, all other regions=0). VMTs have been tested by running the model with the variables for density, the extent of a region's transit network, and an MSA dummy variable. The model was run five times using each MSA dummy variable. None of the regions showed a significant association. Thus, no single region stands out as requiring further investigation to develop variables not yet measured.

The contrast model for automobile ownership was run using the log transformation of density, per capita household income, the distance to nearest transit station, and an MSA dummy variable. The results of the contrast model for automobile ownership only show a significant association for the San Diego MSA. The San Diego automobile ownership contrast model has an adjusted R-squared of 0.466 as opposed to the best-fit model that has an adjusted R-squared of 0.445. This indicates that there is unexplained variation occurring in the San Diego MSA. The fact that a block group is located within the San Diego MSA increases the predictive power of the automobile ownership regression model. Households in the same neighborhood conditions own more automobiles in the San Diego MSA than in the other MSAs evaluated. This occurs for reasons at the regional level in San Diego that remain unexplained.

Another technique to explore the influence of the MSA's regional level variables is to compare fixed effect and mixed effect models. In this approach, variables representing the regions are input into the regression models all at once, rather than region-by-region as with the contrast models.

Further investigation comparing fixed effect and mixed effect regression models for the two dependent variables gives a different indication of the role of regional level variables than the contrast model approach. In these models, the regression for automobile ownership is not statistically better when accounting for regional variation. However, when comparing models for the VMT model, regional effects are found to improve the fit of the regression.

For the VMT regression model, a random effects model with a random slope and intercept for each MSA is better than a fixed effects only model, suggesting both that VMTs differ for each MSA and that each MSA has a different association between VMT and population density at the block group level. In the better fitting models, transit stations per capita are entered as a categorical variable rather than a continuous variable (i.e., only five unique values). This creates a regional identifier similar to using a dummy variable.

The best-fit model for VMTs occurs when an interaction term between the variables for density and extent of regional transit network is used. This gives an adjusted R-squared of 0.504, explaining variance not explained in the initial model that gives an adjusted R-squared of 0.404 (see Table 6). The model shows that an increase in density at the block group scale and an increase in the extent of the regional transit network at the regional scale enhance each other's effects in decreasing VMTs.

The third approach to testing for spatial autocorrelation was to test whether proximity of block groups within regions could explain residual variation. For the VMT regression model, adding a spatial correlation term does not explain any of the residual variance. However, for automobile ownership, proximity of block groups within a region was found to have a significant association.

Spatial relationships of block groups explain a good share of residual variance in the automobile ownership regression model. For the automobile ownership model, including correlation of spatial coordinates, the adjusted R-squared is 0.567 compared with 0.445 in the initial model (see Table 7). This means that the model can best predict

automobile ownership in non-sampled block groups within a given MSA when the distance between these block groups and the sampled block groups is known and taken into account.

The proximity measure in the best fit, spatial model for automobile ownership also competes with the linear hypothesis for the distance to transit variable, rendering it non-significant in the model. Further investigation demonstrated a non-linear relationship between distance to transit and automobile ownership. The spatial model indicates that automobile ownership increases as a function of distance to transit but then gradually levels off and decreases at large distances to stations. There is little theoretical support for this finding, and it deserves further exploration in future studies.

Overall, the tests performed to measure autocorrelation in the block group samples help explain residual variation in both regression models. For VMTs, autocorrelation has no effect within the MSAs. However, it does explain variance in VMTs between the MSAs. Conversely, autocorrelation explains variance in automobile ownership within the MSAs.

4.5 Modifiable Areal Unit Problem

The Modifiable Areal Unit Problem (MAUP) describes the effect that arbitrary areal geometric units have on geographic analyses (Montello & Sutton, 2006). All of the areal units examined in this study are designated by the U.S. Census Bureau and therefore cannot be controlled in study design. Because the block groups used in the sample data vary significantly in size, the MAUP is a potential issue. The potential for the MAUP is

greatest with the block group density variable because it is directly related to the acreage of a block group.

To test for these issues, the largest and least dense block groups in the sample were examined. The three least dense block groups were in the Miami, Portland, and San Diego MSAs. Areal examination revealed that development was generally ex-urban in all three block groups. The vast majority of land was undeveloped in 2000. Small areas of development existed but were mostly spread out (Figure 9).

Figure 9: Largest, Least Dense Block Group in San Diego MSA

If any of these large block groups were to be broken up into smaller units, equivalent to the mean block group size of their MSA, the density variable of the smaller units would most likely have a value of 0. If randomly selected from the smaller units, it would be very unlikely that a unit with any development in it would be selected. As a result, the areal unit has little effect on the density variable throughout most of these block groups. The density value is already extremely low in these block groups and would be only slightly lower if broken up. There are no areas of significant density within any of these block groups.

Chapter 5: Discussion and Conclusion

In general, evidence from the regression models developed in this study support an association between key factors in the built environment and automobile transit behaviors at the neighborhood level. The hypothesis that the built environment influences people to drive less is supported by the research results, however, with a lot of unexplained variation.

The data show that the built environment and the extent of a region's transit network are significantly associated with the daily transportation habits of households. Proximity to a transit station is also significantly associated with the transportation habits of households in urban and suburban areas. Additionally, the number of automobiles owned by a household is significantly associated with the density of the built environment, proximity to a transit station, and household income.

This study put the core principles of smart growth to a robust empirical test. In particular, the study examined the claims that density, street connectivity, and access to jobs, services, and transit would lead to less dependency on automobiles. The findings confirm that density and the extent of a region's transit network are significantly associated with vehicle miles traveled (VMTs). Additionally, automobile ownership is significantly associated with density, transit access, and household income.

These findings have many public policy implications. At the neighborhood level, the findings suggest that density should be encouraged through zoning legislation in areas well served by transit. This echoes the findings of past research that suggests that density is associated with vehicle transportation behavior (Kim & Brownstone, 2010). This is

already a common practice in many jurisdictions through the use of transit-oriented development (TOD). This practice helps to maximize benefits from the large public investment required to build transit lines.

At the regional level, these findings suggest that investments in transit and smart growth projects will have a greater impact on VMTs in areas that have existing transit systems with dense service networks. Because VMTs have a significant association with the extent of a region's transit network, the impact of an expansion to an existing transit system, or organized development around transit, could potentially be greater than that of a new transit line in a region that does not have an existing public transit system. The findings also suggest that the continued use of TOD projects in jurisdictions well served by transit could lead to fewer VMTs. Lastly, the study results suggest that in areas with existing TOD projects, or in areas that are already relatively dense, more density should be considered as part of the master planning process.

Because of the association between household income and automobile ownership, new transit projects could be most effective in low-income areas. Individuals in more affluent neighborhoods own more automobiles and therefore are less likely to use public transit even when the built environment is conducive to driving less. Expansion of transit lines and improved transit access in middle and lower-income neighborhoods could lead to reduced vehicle ownership and VMTs than similar expansions in wealthier neighborhoods. Past studies have also found a significant association between income and vehicle transportation behavior (Shay & Khattak, 2006).

5.1 Future Research

Future research should include additional data to better capture the factors that drive VMTs and automobile ownership. At the time of this study, spatial data for bus routes and stops was inconsistent and therefore not included. Few jurisdictions had publicly available bus stop data. Bus route data were available in some areas but not all. The inclusion of bus data could paint a clearer picture of transit access in future studies, particularly in areas with a limited regional rail transit system.

Almost all of the block group level census data used in the study came from Summary File 3 of the 2000 U.S. Census. At the time of this study, Summary File 3 data from the 2010 U.S. Census had not been released. When these data become available, this study should be reexamined to evaluate the changes in transportation behavior over the ten years from 2000-2010. Longitudinal studies have previously been conducted in an attempt to establish a causal link between the built environment and travel behavior (Cao et al., 2007). A future study could examine the same block groups as this study and report the updated variable data. If the 2010 U.S. Census block groups have changed significantly since the 2000 U.S. Census, a new sample of block groups from the same MSAs could be used to examine changes.

The smallest aggregation at which VMT data were nationally available was the census tract level. Even these data were a one-time estimate based on the MSA level National Household Travel Survey (NHTS) data. For the most part the VMT data that are available at the block group or neighborhood level come from household surveys sponsored by metropolitan planning organizations or universities. These surveys are

almost always for a single MSA. Future studies with the time and funding to do so could gather survey data for VMTs at the block group or neighborhood level across multiple MSAs.

An additional confounding variable that could be used in future research is retail fuel price data. Data for the average retail price of fuel at the MSA level could not be obtained for the year 2000. The cost of fuel could be an important factor in the driving habits of households. These data are available at the MSA level from the mid-2000s to the present from the Oil Price Information Service. Future studies, using 2010 U.S. Census data, could take advantage of MSA-level fuel price data. An additional component of the cost of automobile transportation that should be examined in future studies is the cost of parking. In highly dense areas, such as the downtown area of a major city, parking can be a significant monthly expense for commuters.

Future research should also use more sophisticated and complete measures of walkability. Street connectivity in this study was measured a block group's acreage divided by the number of blocks within the block group. A more sophisticated measure of overall walkability that takes into account infrastructure factors such as sidewalks, elevation change, shade, safety, and quality of walking paths could yield a fuller picture of the walking conditions in a block group. Street connectivity could also be measured using variables related to street connectivity such as intersection density, street network density, and average street block length (Huang et al., 2009).

In future attempts to examine and explain the phenomena explored in this study, additional regions and block groups should be included. This may increase the predictive

power of the models developed. Additionally, it would potentially allow for better analysis of the differences in VMTs and automobile ownership between the MSAs and between regions within MSAs.

This study found that job and retail access are not significantly associated with either VMTs or automobile ownership. These findings are in contrast with some of the tenets of planning theory that suggest that improved access to jobs and shopping will reduce automobile use. At the same time, the methods used in this study to define job and retail centers only capture concentrations of employment and not necessarily areas of mixed-use development. While mixed-use development is a key element of smart growth theory, past studies have shown that mixed land use only slightly decreases overall VMTs (Ingram et al., 2009). Future studies should develop methods for measuring the presence of mixed-use development to see how it affects the significance of job and retail access as they relate to VMTs and automobile ownership.

This analysis could be taken a step further by investigating the link between transportation behaviors and household transportation costs. Once this link is established, it could be used to investigate housing affordability when transportation costs are factored in.

The spatial autocorrelation results suggest that the region a block group is in has an effect on VMTs and automobile ownership. This study is not able to fully explain the variance between regions, and therefore further research into the history and culture of a region could be important. Within regions, the spatial autocorrelation results for automobile ownership indicate that future studies should consider the inclusion of data at

scales between the neighborhood and MSA level. These scales could include sub-regions within MSAs like major road and transit corridors.

Finally, future research should examine the factors that drive VMTs and automobile ownership that were not captured in this study. The best-fit model developed in this study explains approximately 40% of VMTs and 45% of automobile ownership. Additional studies should seek to explain more difficult to understand factors, such as the personal transportation preferences of individuals. Recent research indicates that the built environment may only have a differential impact on walking trips and that an individual's attitude towards walking is more important in shaping walking habits (Joh, Nguyen, & Boarnet, 2011).

A study on the effects of individual preference on transit behavior could be accomplished by examining these variables at an individual scale as opposed to the neighborhood or block group level. Studies could also examine the neighborhood level environment using nested scales. Individual-level data would likely have to be obtained through surveying, and the research design of such a study would need to be altered to ensure that the block group samples collected within each MSA are diverse in the income levels they represent.

Bibliography

Agresti, A., & Finlay, B. (2009). *Statistical methods for the social sciences*. Upper Saddle River, NJ: Prentice Hall.

Allison, P. D. (1999). *Multiple regression*. Thousand Oaks, CA: Pine Forge Press.

Cao, X., Mokhtarian, P., & Handy, S. (2007). Do changes in neighborhood characteristics lead to changes in travel behavior? A structural equations modeling approach. *Transportation*, 34, 535-556.

Center for Neighborhood Technology. (2010, April 2). *Penny wise pound fuelish: New measures of housing + transportation affordability*.

Center for Transit-Oriented Development and Center for Neighborhood Technology. (2006, January). *The affordability index: A new tool for measuring the true affordability of a housing choice.*" Washington, DC: Brookings Institution.

Cervero, R. (1996). Mixed land-uses and commuting: Evidence from the American Housing Survey. *Transportation Research*, 30(5), 361-377.

Cervero, R. (2002). Built environments and mode choice: Toward a normative framework. *Transportation Research Part D*, 7, 265-284.

Cervero, R., & Murakami, J. (2010). Effects of built environments on vehicle miles traveled: Evidence from 370 urbanized areas. *Environment and Planning*, 42, 400-418.

Crane, R. & Crepeau, R. (1998). Does neighborhood design influence travel? A behavioral analysis of travel diary and GIS data. *Transportation Research-D,* 3(4), 225-238.

Gearin, E. (2004). *Up against the sprawl: Public policy and the making of Southern California*. Minneapolis: University of Minnesota Press.

Ewing, R., & Cervero, R. (2001). Travel and the built environment: A synthesis. *Transportation Research Record*, 1780, 87-114.

Gordon, P., & Richardson, H. (1997). Are compact cities a desirable planning goal? *Journal of the American Planning Association*, 63(1), 95-106.

Haas, P., Makarewicz, C., Benedict, A., & Bernstein, S. (2008). Estimating transportation costs by characteristics of neighborhood and household. *Journal of the Transportation Research Board,* No. 2007, 62-70.

Hess, D. B., & Ong, P. M. (2002). Traditional neighborhoods and automobile ownership. *Transportation Research Record*, Paper No. 02-2594, 35+.

Hu, P., Reuscher, T., Schmoyer, R., Chin, S. (2007, May 2). *Transferring 2001 National Household Travel Survey*. Oak Ridge National Laboratory.

Huang, L., Stinchcomb, D., Pickle, L., Dill, J., & Berrigan, D.. (2009). Identifying clusters of active transportation using spatial scan statistics.   *American Journal of Preventive Medicine*   37(2).

Ingram, G., Carbonell, A., Hong,Y., & Flint, A., eds. (2009**).** *Smart growth policies: An evaluation of programs and outcomes*. Lincoln Institute of Land Policy.

Kim, J., & Brownstone, D. (2010). *The impact of residential density on vehicle usage and fuel consumption*. University of California: Department of Economics.

Joh, K., Nguyen, N., & Boarnet, M. (2011). Can built and social environmental factors encourage walking among individuals with negative walking attitudes? *Journal of Planning Education and Research* from http://jpe.sagepub.com/content/early/2011/11/29/0739456X11427914.

Krizek, Kevin J. (2003 Summer). Residential relocation and changes in urban travel: Does neighborhood-scale urban form matter? *Journal of the American Planning Association,* 69(3), 265.

Montello, D., & Sutton, P. (2006). *An introduction to scientific research methods in geography*. Thousand Oaks: Sage Publications.

O'Sullivan, D., & Unwin, D. (2010). *Geographic information analysis*. Hoboken, NJ: John Wiley & Sons.

Shay, E., & Khattak, A. (2006). Autos, trips and neighborhood type: Comparing environmental measures. *Transportation Research Record,* Paper 07-1119, 1-18.

United States. Office of Management and Budget. (2000, December 27). Standards for defining metropolitan and micropolitan statistical areas. *Federal Register,* 65, 249.

United States Census 2000. *Summary File 3 (SF 3),* from
    http://www.census.gov/census2000/sumfile3.html.

Zhang, M. (2006). Travel choice with no alternative: Can land use reduce
    automobile dependence*? Journal of Planning Education and Research*, 25,
    311-326.