Spring 2013

The Role of Proximity in Reducing Auto Travel: Using VMT to Identify Key Locations for Development, from Downtown to the Exurbs

Robert B. Case
Old Dominion University, rcase@hrtpo.org

THE ROLE OF PROXIMITY IN

REDUCING AUTO TRAVEL

USING VMT TO IDENTIFY KEY LOCATIONS FOR DEVELOPMENT,

FROM DOWNTOWN TO THE EXURBS

by

Robert B. Case, PE
B.A. May 1982, University of Virginia
B.S. December 1987, Old Dominion University
M.S. June 1993, Georgia Institute of Technology

A Dissertation Submitted to the Faculty of
Old Dominion University in Partial Fulfillment of the
Requirements for the Degree of

DOCTOR OF PHILOSOPHY

CIVIL ENGINEERING

OLD DOMINION UNIVERSITY
April 2013

Approved by:

____
Asad Khattak (Director)

____
Mecit Cetin (Member)

____
ManWo Ng (Member)

ABSTRACT

THE ROLE OF PROXIMITY IN
REDUCING AUTO TRAVEL

USING VMT TO IDENTIFY KEY LOCATIONS FOR DEVELOPMENT,
FROM DOWNTOWN TO THE EXURBS

Robert B. Case
Old Dominion University, 2013
Director: Dr. Asad Khattak

The purpose of this dissertation is to discover the VMT impact of each level of proximity in order to help government identify key locations for housing development, and thereby lower VMT and reduce dependence on foreign oil. By discovering the VMT impact of each level of proximity, this dissertation provides a) the first known means of calculating the proximity-based VMT benefit of subject locations by individual proximity level, and b) the new finding that it is *likely* that high VMT benefit can be achieved at moderate proximity levels acceptable to many households, enabling representative governments to be politically successful while promoting housing in locations that will lower the average VMT of the population.

After discussing the impetus for the work, this dissertation presents a theory of the determinants of VMT, searches the literature for appropriate techniques for empirical analysis of the proximity-VMT relationship, and presents results of the empirical research to be expected based on the presented theory and literature.

Empirical efforts are used to discover VMT impact by proximity level using three differing measures of proximity: density, distance-threshold-based total opportunities,

and centrality. In the first effort, national data is used to discover VMT impact by proximity level, for both population and employment density. In order to determine the role played by alternative modes in the VMT-density curves of the first effort, the second effort uses national data to discover the impact of each level of density on usage of alternative modes. In the third and final effort, data from Hampton Roads, Virginia, are used to discover the VMT impact of each level of opportunity and centrality.

Governments can apply the discovered VMT impact of each level of proximity—via a described "VMT Benefit Technique"—to accurately determine the VMT benefit of a given location, and use the VMT benefits of a set of candidate areas to select key locations for development.

In addition, the discovered VMT impact of each level of proximity informs the key hypothesis of this dissertation that there exists a sweet spot on the VMT-proximity curve that has high VMT benefit and a proximity level acceptable to many households. Although the hypothesis tests indicate that it is *not certain* that the sweet spot exists, the mean coefficients of the models indicate that it is *likely* that the sweet spot exists, i.e. that there are high-VMT-benefit proximity levels acceptable to many households. The overall implication of this is that representative governments in the U.S. who promote housing development at these moderate levels of proximity will not only lower average VMT in the short term, but they will not be punished politically for doing so, and therefore may be successful in thereby lowering average VMT in the long term.

In summary, the dissertation provides encouragement to governments hoping to lower average VMT and an accurate method of calculating VMT for choosing SGAs with which to actually lower average VMT. It is hoped that this combination will help U.S. governments become independent of foreign oil.

DEDICATION

This dissertation for using VMT to identify key locations for housing development to reduce dependence on foreign oil is dedicated to God.

ACKNOWLEDGEMENTS

Research guidance was provided by Dr. Asad Khattak of Old Dominion University (ODU). Dr. Mecit Cetin and Dr. ManWo Ng served on the dissertation committee. Geocoded NHTS data was obtained through the help of Paul Agnello and Ju-Yin Chen of the Virginia Department of Transportation (VDOT). Susan Liss (FHWA contractor) provided NHTS guidance. Virginia Employment Commission (VEC) data was obtained through the help of Dale Stith and Benito Perez of the Hampton Roads Transportation Planning Organization (HRTPO). Dale Stith (HRTPO) provided zone-to-zone distance and travel time data for Hampton Roads. Some of the funding for this degree came from ODU and the HRTPO.

Dr. Michael Meyer taught me how to think about transportation. John Crosby helped inspire my love of geography. Dr. Camelia Ravanbahkt and Dale Castellow identified me as a researcher. Dwight Farmer challenged me to learn to be a researcher. My wife, Bobbi J. Case gave me moral support and editing suggestions.

TABLE OF CONTENTS

Section Page

LIST OF TABLES vii

LIST OF FIGURES ix

Chapter

I: INTRODUCTION 1

II: IMPETUS FOR THIS DISSERTATION 6

III: PREPARATION FOR EMPIRICAL ANALYSIS 19

IV: EMPIRICAL ANALYSIS 51

Effort #1: Identifying Key Proximity Levels for VMT Using National Data 52

Effort #2: Identifying Key Proximity Levels for Alternative Modes Using National Data 86

Effort #3: Identifying Key Proximity Levels for VMT Using Hampton Roads Data 108

VMT vs. Opportunity 120

VMT vs. Centrality 134

V: CONCLUSION 144

REFERENCES 151

VITA 155

LIST OF TABLES

Table Page

1. VMT-Related Issues in the Transportation Literature 8

2. Summary of Theorized Determinants of Annual Household VMT 21

3. Comparison of ANNMILES to FHWA's Highway Statistics 53

4. Similarity Between Full Dataset and Analysis Dataset 55

5. Summary of Theorized Determinants of Annual Household VMT 56

6. Descriptive Statistics 60

7. VMT-Density OLS Regression Results 65

8. Prevalence of Employment Density Levels in the U.S. 73

9. Hypothesis Testing Worksheet based on Employment Density Curve 75

10. Prevalence of Population Density Levels in the U.S. 81

11. Hypothesis Testing Worksheet based on Population Density Curve 83

12. Similarity Between Full Dataset and Analysis Dataset 88

13. Summary of Theorized Determinants of Annual Household VMT 89

14. Descriptive Statistics- Modal Detail 92

15. Descriptive Statistics 93

16. Alternative-Mode-vs.-Density Logistic Regression Results 96

17. Prevalence of Employment Density Levels in the U.S. 102

18. Prevalence of Population Density Levels in the U.S. 106

19. Similarity Between Full Dataset and Analysis Dataset 109

20. Summary of Theorized Determinants of Annual Household VMT 110

21. Descriptive Statistics 117

22. Initial VMT-Opportunity OLS Regression Results 119-121

23. Final VMT-Opportunity OLS Regression Results 124

24. Prevalence of Opportunity Levels in Hampton Roads 130

25. Hypothesis Testing Worksheet based on VMT-Opportunity Curve..........................132

26. VMT-Centrality OLS Regression Results 135

27. VMT Benefit at Each Level of Centrality..138

28. Prevalence of Population Centrality Levels in Hampton Roads...............................141

29. Hypothesis Testing Worksheet based on VMT-Centrality Curve 142

LIST OF FIGURES

Figure Page

1. Energy Dependency 9
2. Travel Behavior by Population Density, U.S. Total. 14
3. Likelihood of Owning One, Two, or Three Vehicles Relative to Owning Zero Vehicles, by Population Density. 15
4. Expected Shape of VMT-Proximity Curve 26
5. Expected Sweet Spot on VMT-Proximity Curve 27
6. Key Relationships- Effort #1 59
7. Histogram for Dependent Variable (ANNMILES) 62
8. VMT vs. Employment Density of Household Census Tract 72
9. VMT vs. Employment Density of Household Census Tract 73
10. VMT vs. Population Density of Household Census Tract 77
11. VMT vs. Population Density of Household Census Tract 80
12. Usage of Alternative Modes for Work vs. Employment Density of Household Census Tract 99
13. Usage of Alternative Modes for Work vs. Population Density of Household Census Tract. 102
14. Hampton Roads, Showing Centers 111
15. Key Relationships- Effort #3 113
16. VMT vs. Opportunity 127
17. VMT vs. Opportunity 128
18. VMT vs. Centrality. 137

19. Hampton Roads Employment Locations 138

20. VMT vs. Centrality 139

CHAPTER I

INTRODUCTION

Preview, Purpose, and Research Objective

Although auto travel provides great benefits to the individual—enabling the traveler to quickly participate in a variety of desirable activities which occur away from the home—it also causes unintended consequences to American society, including environmental threats, roadway congestion, and demand for energy resulting in a world-wide battle for petroleum resources. Since these costs or disbenefits are not directly borne by the auto driver, they are not kept in check by market forces. There is incentive therefore for the representative governments in the U.S., among other responses, to reduce the amount of auto travel while maintaining individual activity. *This dissertation therefore seeks to help governments reduce auto travel.*

Given that much of U.S. electric power is generated with coal and nuclear fuel, auto travel produces disbenefits regardless of auto engine type. *This dissertation therefore uses vehicle miles traveled (VMT) to measure auto travel.*

Given that the cost of the aforementioned world-wide battle for petroleum resources is approximately one trillion dollars, *this dissertation is dedicated to energy independence.*

The amount of VMT conducted by a household is a complex function of 1) the nature of the household (wealth, family structure, culture, etc.), 2) its economic environment (energy supply, energy taxes, tolls, subsidies, etc.), and 3) its physical environment, i.e. a) the location of origins/destinations, and b) the transportation infrastructure between them. *Of these three, this dissertation will focus on changing the physical environment to reduce VMT.*

Although proposals of changes to the physical environment for lowering VMT have many different names—e.g. mixed-use developments, infill, higher density, downtown redevelopment, transit-oriented design, smart growth, public transit, and traditional neighborhood development—the vast majority of these proposals are promoted because they make origins and destinations proximate and/or they supply the infrastructure for alternative modes that require proximity. Given that proximity reduces auto travel distances *and* provides environments in which government is willing to invest in higher-speed alternative modes (bus, rail) which compete with auto travel better than lower-speed alternative modes (walking, biking), *this dissertation will focus on using proximity to reduce VMT.*

In the ideal mono-centric circular metro where density decreases with distance from center, as centrality decreases, both neighborhood-based proximity and regionally-based proximity decreases. Because most metros resemble—albeit imperfectly—this ideal metro, centrality is a proxy for proximity. Because centrality is a proxy for proximity (which reduces auto travel) and centrality is easier to picture and measure than true proximity, centrality—achieved by locating new housing near the center, via, for example, downtown redevelopment or urban growth boundaries—has also been promoted as a way to lower auto travel. *Therefore, to accompany its focus on proximity, this dissertation will also focus on using centrality to reduce VMT.*

The literature indicates that proximity reduces auto travel, the latter often measured by vehicle miles of travel (VMT). Some governments have therefore used their financial and regulatory powers to promote central living. But how central should new housing be? Should it all be downtown? What about the suburbs? Is their VMT impact or "VMT signature" more similar to that of downtown residences or that of exurban residences? In other words, what is the shape of the VMT-centrality curve? And, similarly, what is the shape of the VMT-proximity curve? If these curves have a curved shape, where are the bends in the curves? For example, how much centrality or proximity is "enough", beyond which little benefit is realized? As shown in the Impetus section below, the literature does not provide the VMT benefit by proximity level, creating a gap in knowledge. Therefore, *the purpose of this dissertation is as follows*:

> to discover the VMT impact of each level of proximity in order to help government identify key locations for housing development, and thereby lower VMT and reduce dependence on foreign oil.

From this purpose, *the research objective is as follows*:

> to discover the VMT impact of each level of proximity.

In order to discover the shape of these VMT curves, this dissertation will explore the proximity-VMT relationships theoretically, then empirically. The empirical analysis will use rigorous statistical methods to explore the travel of households living in various environments across the U.S. and across one large metro area, Hampton Roads. It will regress VMT on certain proximity measures—and control variables—using, therefore, a survey containing travel data, socio-economic data, and location data—the 2009 NHTS—from which centrality and proximity to destinations can be measured.

By measuring the travel impact of various levels of proximity, this dissertation will inform governments—the target audience of this dissertation—on the relative benefits of promoting housing construction/renovation at various levels of proximity and centrality—from downtown, to inner-suburb, to outer-suburb, to exurbs.

Definitions- Proximity, Centrality, Opportunity, and Accessibility

In this dissertation, "proximity" will be used to refer to the physical closeness of origins and destinations. When referring to the proximity of a given household, the term will be used as an attribute of that household's location that is the degree to which the home is located *near* activity destinations, e.g. schools, places of work, shopping centers, friends' homes. Therefore, a household's proximity is not based on the speed of the transportation systems in its environment. A household's proximity is a function solely of the number and type of destinations near the home and the travel distances to them.

In this dissertation, "centrality" will be used as an attribute of a home's location that is the degree to which the home is located near the metro's *center*. Locations at the center have maximum centrality; locations at the metro edge have minimum centrality. Although, as discussed above, centrality is a proxy for proximity, for the sake of convenience, centrality is also discussed herein as a measure of proximity.

In this transportation dissertation, an accessible destination is considered an "opportunity." Therefore, "total opportunities" is the sum of accessible destinations, and "distance-threshold-based total opportunities" is the sum of destinations within a given distance of the location of the traveler. "Distance-threshold-based total opportunities" is one of the measures of a household's proximity used in this dissertation. Note also that the singular term "opportunity", when used as an attribute of a home's location, is shorthand for the sum of destinations within a given distance.

Finally, the term “accessibility” must be defined for this dissertation. In preparation for doing so, definitions from the literature are examined. According to Chen et al. (*22*), accessibility is “…the ease (or difficulty) with which activity opportunities may be reached from a given location using one or more modes of transportation.” This “ease” includes time and distance. According to Handy (*27*), “Accessibility, as generally defined, consists of two parts: a transportation element or resistance factor and an activity element or motivation factor…” Likewise this “resistance factor” includes time and distance. The accessibility definition in TRB’s *Highway Capacity Manual* (HCM), however, is more restrictive, covering only time: “The percentage of the populace able to complete a selected trip within a specific time.” Therefore, whereas proximity and centrality have no time component, accessibility—whether comprehensively or restrictively defined—always has a time component.

In this dissertation the simple term "accessibility"—in accordance with the less restrictive definitions in the literature—will refer to the degree to which residents of the subject home can *easily* access desirable activity destinations. Accessibility, therefore, will be considered a function of 1) the number of local activity destinations (by type), and 2) the ease of reaching them, e.g. the “costs” (e.g. distance, time, fares) of the modal paths (e.g. sidewalks, bus routes, and roads) that connect the household to those locations. When time is the only portion of that cost considered—as in the case of the above HCM definition—this dissertation will use the term “time-based accessibility.”

Other Definitions

Finally, in this dissertation:

- the term “metro” is used to refer to a metropolitan area

- the term “neighborhood-based proximity” is used to refer to the degree to which there are destinations in the near vicinity of the subject household

- the term “regionally-based proximity” is used to refer to the degree to which the subject household is near to destinations in the region in which it lies

CHAPTER II

IMPETUS FOR THIS DISSERTATION

Impetus for Reducing VMT

Impetus for Reducing VMT in the Literature

According to the literature, a) auto travel causes a variety of disbenefits, and b) vehicle miles of travel (VMT) is often chosen to represent auto travel in efforts undertaken to reduce auto travel. According to Cervero and Murakami (*6*), "VMT per capita is widely viewed as the strongest correlate of environmental degradation and resource consumption in the transportation sector—as individuals log more and more miles in motorized vehicles, the amount of local pollution (eg particulate matter) and global pollution (eg greenhouse gas, or GHG, emissions) increases, as does the consumption of fossil fuels, open space, and other increasingly scarce resources." Dunphy and Fisher (*16*) focused on disbenefits related to air quality, and they chose VMT as their auto travel disbenefits measure: "Vehicle miles of travel (VMT) has joined vehicle trips as a critical travel demand indicator because it correlates closely to air quality." Schimek (*13*) added "congestion" to the list of auto travel by-products: "The source of the interest in travel behavior has been concern for the air quality, congestion, and quality-of-life impacts from increasing automobile usage." Ortuzar (*14*) added crashes to the list: "…*side-effects* [original emphasis] associated with the production of transport services: accidents, pollution and environmental degradation in general." Salon et al. (*25*) added human health and social interaction as VMT issues: "These [benefits of reduced VMT] include alleviating traffic congestion, reducing air pollution, reducing greenhouse gas emissions,

…improving public health through increased exercise, and enhancing interactions within our communities." Shay and Khattak (*33*) add economic considerations ("public and household budgets") to the list of negative effects.

TABLE 1 VMT-Related Issues in the Transportation Literature

Article/Book (*endnote*)	Congestion	Pollution, GHG	Oil Consumption, Oil Importation, Security	Open Space Consumption	Quality of Life, Health, Social Interaction, Finances	Crashes, Safety
Cervero & Murakami (*6*)		√	√	√		
Dunphy & Fisher (*16*)		√				
Schimek (*13*)	√	√			√	
Ortuzar (*14*)		√				√
Salon et al. (*25*)	√	√	√		√	
Shay & Khattak (*33*)		√	√	√	√	√
TRB Special Report 299 (*15*)		√	√			

Tables.xlsx

The Problem of Depending on Foreign Oil in the Literature Several transportation authors consider oil importation as a disbenefit of auto travel. According to Salon et al. (*25*): "These [benefits of reduced VMT] include…reducing our dependence on foreign oil…" Apparently alluding to the global effects of U.S. oil importation, Shay and Khattak (*33*) included "security" in their list of problematic issues related to auto travel, calling auto-based travel "untenable" for this (and various other) reasons. According to unpublished slides by Khattak, the specific problems created by oil importation include:

> "I. Supply is subject to embargo (e.g. 1973 Arab Oil Embargo)
>
> II. Cost is subject to shocks (e.g. 1990 Iraq invasion of Kuwait decrease oil supply by more than 4 million barrels per day)"

The authors of TRB Special Report 299 (*15*) also pointed to the problem of importing oil for auto travel: "[Transportation]…consumes twice as much petroleum as the United States produces annually."

Oil importation is, of course, also a concern of people outside of transportation academia. According to Neela Banerjee of the McClatchy-Tribune Information Services, “ever since the Nixon administration, every president has pledged to reduce the United States’ dependence on foreign oil…” (*36*) Unfortunately, “The U.S. imported 45 percent of its petroleum last year…” (*39*)

FIGURE 1 Energy Dependency. (*37*)

Impetus for Reducing VMT in this Dissertation

Because 1) this dissertation seeks to inform government action (as stated above), and 2) U.S. government has a representative form, the impetus for reducing VMT in this dissertation is largely the desires of the voting public concerning VMT. It is assumed in this dissertation that the voting public seeks to reduce many of the auto travel disbenefits listed in the literature concerning limiting VMT, i.e. congestion, pollution, oil importation, and crashes. Of these—given that the cost of the aforementioned world-wide battle for petroleum resources is approximately one trillion dollars—this dissertation is dedicated to reducing VMT in order to reduce oil importation.

The Problem of Depending on Foreign Oil in this Dissertation Importing such a large amount of petroleum has caused and will cause significant problems for Americans. We have experienced the impact of instant reductions in the availability of oil overseas. In 1973, the U.S. suffered economically from the Arab Oil Embargo. But more importantly, we have experienced—and are experiencing—the impact of U.S. military involvement in the Persian Gulf , a response to the natural insecurity of depending on foreign nations for a key commodity such as oil. A shift in the balance of power in the Persian Gulf, i.e. Iraq's takeover of Kuwait, led to Gulf War I. Placing troops and planes in Saudi Arabia (the site of Muslim holy places) led—in part—to al Qaeda's 9-11-01 attacks. Gulf War I and 9-11 led to the current wars in Iraq and Afghanistan, which has cost the U.S. (to date) thousands of lives and approximately one trillion dollars.

It should be noted that, at the end of 2011, the media outlets published many positive analyses in anticipation that 2011 would be the first year since 1949 in which the U.S. would be a net exporter of petroleum products, i.e. exporting more petroleum than it imports. Titles included "The Coming Day of Energy Independence" and "Foreign Oil? Who Needs It!" Unfortunately, in contrast to these misguided headlines, the fact that the U.S. is now a net oil exporter has not eliminated the problems caused by massive oil importation. Because the global oil market—like any market—is largely driven by price, it was cheaper in 2011 for the U.S. to buy 45% of its oil overseas than to buy this 45% from domestic producers. Likewise, producers got a higher net price for that portion of domestic production which was sold overseas in 2011 than they would have gotten had they sold it domestically. Therefore, disturbances overseas forcing U.S. consumers to buy some of that 45% from domestic production—although not necessarily causing shortages—would cause increased prices for U.S. consumers. This price threat may explain why the U.S. is still fighting in Iraq and Afghanistan, recently added Libya to the list, and may soon add Syria and Iran. As the Associated Press stated, "the United States is nowhere close to energy independence" (*38*).

Consequently, it is hoped that this dissertation will help governments lower U.S. VMT to lower oil importation and thereby reduce the incentive for U.S. military intervention around the world.

Using Proximity to Reduce VMT

Using Proximity to Reduce VMT- Popular Proposals

In the planning departments of government, and in the supporting field of transportation research, there are several common proposals designed, at least in part, to reduce VMT, as follows:

- mixed-use developments
- infill
- higher density
- downtown redevelopment
- transit-oriented design
- smart growth
- public transit, and
- traditional neighborhood development

The common component of these proposals is that they make origins and destinations proximate and/or they supply the infrastructure for alternative modes (walking, biking, bus, rail) that require proximity.

Using Proximity to Reduce VMT- this Dissertation

Given that 1) many common proposals for lowering VMT—as listed above—are based on proximity, 2) theory leads one to expect that proximity reduces VMT, and 3) the literature indicates that proximity reduces VMT—the latter two points to be discussed in the Preparation section below—this dissertation will focus on using proximity to reduce VMT.

This dissertation will also explore using centrality (i.e. placing homes near the metro center) to reduce VMT. In the ideal mono-centric circular metro where density decreases with distance from center, as centrality decreases, both neighborhood-based proximity and regionally-based proximity decreases. Because most metros resemble—albeit imperfectly—this ideal metro, centrality is a proxy for proximity. Because centrality is a proxy for proximity (which reduces auto travel) and centrality is easier to picture and measure than true proximity, this dissertation will also examine using centrality to reduce VMT.

Identifying Key Locations for Development: The Need for Research to Estimate the VMT Impact of Each Level of Proximity

The lack of VMT impact by proximity level in the literature and the value of knowing VMT impact by proximity level establishes a need for research to determine VMT impact by proximity level.

The Lack of VMT Impact by Proximity Level in the Literature

Unfortunately, the literature does not provide government with an understanding of the VMT signature of each level of proximity. Although many of the studies which explored the proximity-VMT relationship examined the existence, strength, and slope of that relationship, it appears that most did not delve into the *shape* of the relationship. For example, in their VMT models, Bento et al (*18*), Cervero and Duncan (*3*), Cervero and Murakami (*6*), and Kockelman (*19*) simply reported the coefficients of their proximity variables, thereby treating the VMT-proximity curve as one of constant slope. Likewise:

- In 2005, Golob and Brownstone found: “Comparing two households that are similar in all respects except residential density, a lower density of 1,000 housing units per square mile [block group measure] implies a positive difference of almost 1,200 miles per year…” (*21*)

- In 2007, Ewing et al. reportedly found "a 0.152 percent reduction in VMT from a 1 percent increase in population density on the basis of their longitudinal model…" (*1*)

That portion of the literature which did address the shape of the proximity-VMT relationship did so rudimentarily. Using zip code density to measure proximity, Dunphy and Fisher (*16*) identified one bend in the VMT-density curve at 4,500 persons per square mile (zip code measure):

> “National data suggest that even doubling density [of zip codes] from the lowest levels typical in a low-density suburb has little effect on reducing travel. Above this level, higher densities begin to have a significant impact on driving, with each doubling of residential density resulting in an approximate reduction of 10 to 15 percent in per capita driving.”

The Dunphy and Fisher curve (with VMT on the right-hand axis) is shown below:

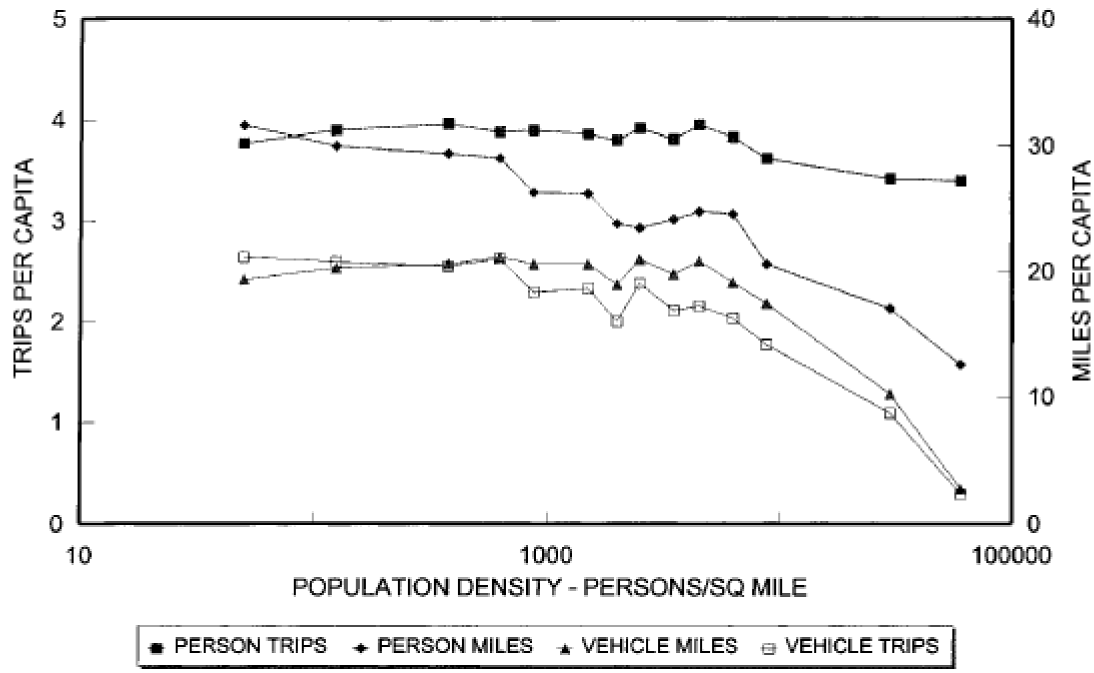


FIGURE 2 Travel Behavior by Population Density, U.S. Total.

Note that it is impossible to isolate the impact of density from this research because Dunphy and Fisher did not control for socio-economics (e.g. income) in producing it.

Schimek (*13*), who also examined the travel-density relationship, postulated—but did not investigate—a bend in the curve:

> "Because all three of these [density] effects—better walking, better transit, and more expensive car use— occur simultaneously, the overall effect of density may be nonlinear. There may be a threshold above which these factors begin to have a strong effect on travel behavior."

Concerning the relationship between density and vehicle ownership (a key determinant of VMT), Dunphy and Fisher (*16*) found one bend in the curve at 4,500 persons per square mile (zip code measure), above which vehicle ownership declines more rapidly. Likewise, Walls, Harrington, and Krupnick (*20*) found a bend in the curve at 4,500 persons / sq. mi. (zip code measure) above which ownership declines rapidly.

And they found a second bend at 10,000, above which vehicle ownership declines slowly. The S-shaped Walls curve is shown below:

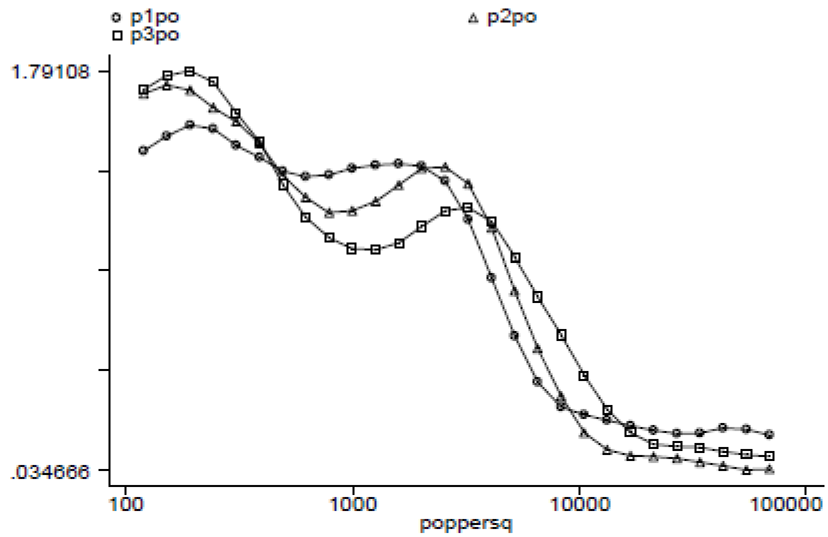


FIGURE 3 Likelihood of Owning One, Two, or Three Vehicles Relative to Owning Zero Vehicles, by Population Density.

Because these two analyses 1) measured auto ownership (as opposed to VMT), and 2) did not control for socioeconomics (e.g. income), instead of identifying the density levels at which the VMT-density curve bends, they merely suggest that the VMT-density curve has one or more bends.

Finally, two papers presented travel-vs.-proximity coefficients/elasticities at various levels of proximity, but did not report expected VMT benefit at those various proximities. Yoon, Golob, and Goulias (*26*) divided their dependent variable (solo driving time) into deciles and divided their land use independent variables (e.g. household density, retail employees within 10km) into deciles and used an Ordered Logit Regression to calculate coefficients for each land use decile. Interpreting these coefficients, the authors simply stated that "households located in areas with lower local (within 10 km) retail accessibility spend more time solo driving than households located in the highest level of retail accessibility", i.e. the authors did not 1) plot a curve, 2) discuss the shape of the curve, or 3) present expected values of solo driving time for each proximity category, or—as one might do for an ordered logistic regression—provide odds ratios or chances for falling into each of the driving time deciles. In summary, even though they stated that "spatial variables can contribute nonlinear and even non-ordinal effects", they did not investigate those effects.

The second paper found to measure VMT-vs.-proximity coefficients/elasticities at various levels of proximity was written by Boarnet et al. (*29*). In that paper, the authors analyzed Los Angeles data subdivided by quintiles of proximity using a "stratified sample" approach and a "spline regression" approach. Although the authors indicated that the VMT-proximity curve is non-linear, and they reported elasticities (VMT vs. proximity) for various proximity ranges, they did not report expected VMT benefit at various proximities.

In summary, the existing transportation literature does not provide VMT impact by individual proximity level. For comparing the VMT impact of certain proximity levels to each other, the literature only provides coefficients of the slope of proximity-VMT relationship.

The Value of VMT Impact by Proximity Level

Government can use an understanding of the relationship between individual proximity levels and VMT as one input in the process of identifying key locations for development. One component of comprehensive planning is identifying key locations for development (i.e. areas in which government would prefer development occur), often referred to as "strategic growth areas" (SGAs). Whereas government currently considers many non-VMT factors when choosing these areas—e.g. availability of land for development or redevelopment, existing supportive infrastructure, etc.—if it had a refined method of estimating the expected VMT impact of the proximity of the location of candidate SGAs, it could add VMT reduction as a factor in the process of identifying key locations for development. Government could use the VMT impact of each level of proximity to score candidate SGAs on expected VMT impact, and combine those scores with other considerations (land availability, infrastructure, etc.) to select the best areas for development. Once these areas have been identified, government could use its regulatory powers (e.g. zoning, use permits) and financial resources (e.g. provision of public works [schools, roads, utilities, and parks] which attract/enable development) to promote housing development in those areas and thereby reap the related VMT impact.

Based only on common sense, some analysts in government currently understand that new households with high proximity tend to produce less VMT than those with low proximity, but they do not know how much proximity is necessary to provide a desired VMT benefit. Those analysts with knowledge of the slope of the proximity-VMT relationship from the above-reviewed literature have more than a common sense understanding of the proximity-VMT relationship, but given—as will be shown—that the true relationship between VMT and proximity is not linear, any calculation they may make (using these slopes/coefficients) of the proximity necessary for a desired VMT benefit will be inaccurate.

Furthermore, given that only the slope of the relationship can be found in the existing transportation literature, even the informed analyst is currently forced to assume that low proximity provides low VMT impact, medium proximity provides medium VMT impact and high proximity provides high VMT impact. Therefore, if the analyst's government desires to lower average VMT, he/she seeks high VMT impact which—according to the current literature—can only be achieved with high proximity. (By refining the measurement of the relationship between proximity and VMT—by measuring the VMT impact at each proximity level—this dissertation will reveal that *high* VMT impact can be achieved with *moderate* proximity.) But because high proximity areas, being located near metro centers, often have lot sizes and school quality unacceptable to many persons, promoting housing development only in high proximity areas will fail. High proximity housing will become partially empty, lower proximity housing will become scarce, and those politicians which promoted this occurrence will be replaced. Consequently, given the representative form of American governments, knowing the VMT benefit of each level of proximity is critical to the success of governments using development to lower VMT.

Summary of the Need for Research

Given, as shown above, 1) that the literature does not provide government with an understanding of the relationship between proximity and VMT at various levels of proximity, and 2) that it is necessary for government to know the VMT impact of each level of proximity in order to a) accurately estimate the VMT impact of candidate SGAs, and b) successfully use proximity to lower VMT (i.e. in a manner amenable to the voting public), there is need for a means of estimating the VMT benefit at each level of proximity, the primary original work and product of this dissertation.

CHAPTER III

PREPARATION FOR EMPIRICAL ANALYSIS-

THEORY, TECHNIQUES, EXPECTED RESULTS, AND HYPOTHESES

Annual VMT Theory Overview

The theoretical determinants of annual household VMT can be identified by examining human nature. It is assumed that people, due to their nature, desire personal interaction, recreation, productivity, rest, and consumption of goods. Because these activities often occur outside the home, travel is desirable. Given the constraint of the 24-hour day and limited income, people seek to minimize the amount of time and money spent on traveling. Because auto travel is generally quicker but more expensive than alternative modes, *it is expected that a) household income, and b) public transit service level are determinants of mode choice and therefore of annual household VMT.* In addition, it is assumed that, to a certain degree, people have individual biases toward the various travel modes, and therefore choose where they live, in part, in accordance with those biases (known in the literature as "self-selection"). Consequently, *it is expected that modal biases are a determinant of mode choice and thereby annual household VMT.*

Given the high incomes in the United States, auto travel is the most common mode choice of Americans. Because laws limit driving to persons who have reached a certain age, *it is expected that the age of persons is a determinant of annual household VMT.* Because high roadway speeds allow drivers to reach distant but desirable destinations without spending much time, *it is expected that time-based accessibility is a determinant of annual household VMT.*

Given the popularity of driving, that which affects the amount of travel also generally affects the amount of *auto* travel or VMT. Due to the constraint of the 24-hour day and limited income, travel—and therefore auto travel—is a function of time and money issues. Because short trips generally save time and money, *it is expected that the proximity of a household to destinations is a determinant of annual household VMT.* Likewise, centrality being a proxy for proximity, *it is expected that the centrality of a household is a determinant of annual household VMT.* Given that internet connectivity allows persons to achieve activity without the time or expense of traveling, *it is expected that internet connectivity is a determinant of annual household VMT.* Because the things and activities we desire often cost money, a household needs money, which comes either through payments (e.g. retirement income, welfare) or work, the latter usually occurring outside the home. *It is expected therefore that work status is a determinant of annual household VMT.*

Time and money, however, are not the only things that affect travel and thereby auto travel. Given that persons are the entities which have the aforementioned desires which induce travel, *it is expected that the number of persons in a household is a determinant of annual household VMT.* But all persons do not have equal desire and ability to travel. Due to the nature of men and women, women are typically more oriented toward the home than men. Therefore, *it is expected that gender is a determinant of annual household VMT.* Given that healthy/whole people are better able to travel, *it is expected that disabilities are a determinant of annual household VMT.*

TABLE 2 Summary of Theorized Determinants of Annual Household VMT

Determinant	Universe
Proximity	Household
Internet Connectivity	Person, Household
Time-Based Accessibility	Household
Public Transit Service Level	Household
Travel Mode Biases ("self-selection")	Person
Socio-economics	
Work Status	Person
Income	Person, Household
Gender	Person
Age	Person
Number of Persons	Household
Disabilities	Person

Tables.xlsx

It should be noted that 1) some of these VMT determinants affect VMT through the intermediate step of auto ownership (as discussed the Auto Ownership section below), and 2) some VMT determinants affect other VMT determinants. However, because the impact of proximity and socio-economics on transit infrastructure occurs over time, this impact is not applicable to the cross-sectional models of this dissertation.

In the following sections, some of the determinants of VMT summarized in the table above are examined in depth concerning theory, measurement, expected results, and hypotheses; starting first with “Proximity”, the variable of interest to this dissertation. Secondly “Socio-economics” are investigated, followed by a look at “Auto Ownership”, a step between proximity, transit infrastructure, and socio-economics (on one hand) and VMT (on the other). Finally, “Time-Based Accessibility” and “Travel Mode Biases” are examined, followed by a look at “Subsets of VMT.”

Proximity (and related hypotheses)

This dissertation's detailed examination of the determinants of annual household VMT begins with proximity, its variable of interest.

Theory, Conceptual Structure, and Hypotheses of Proximity's Impact on VMT

Whereas the postulate that proximity's impact on VMT is generally based on the desire to minimize travel time and cost was presented briefly in the overview above, it will be theorized in detail below. The theoretical impact of proximity on VMT (i.e. *distance* traveled in *auto*) will be examined by looking at the two components of VMT:

1) the choice of the *auto* mode, and

2) given the choice of auto, the *distance* traveled to destinations of activity.

Proximity's Impact on Choosing Auto First, concerning the choice of mode, proximity reduces reliance on the auto as follows. Comparing the auto to alternative modes (walk, bike, public transit), the auto is generally faster, but the auto is pricier. Therefore, 1) there is an income line above which auto is generally used and below which alternatives are generally used, and 2) the income line (and therefore mode choice) shifts with changes in the difference in price (between auto and alternative modes), and changes in the travel time difference (between auto and alternative modes). Thus, proximity affects mode choice via affecting price and travel time.

In regards to price, places with high proximity tend to have high neighborhood-based density, as discussed above. In areas with high density, land is naturally more valuable, and a price is often charged for parking autos. Given basic economics then, proximity reduces the occurrence of choosing the auto mode by establishing (or increasing) a price for parking autos.

In regards to travel time, proximity reduces the choosing of auto by reducing the travel time difference between auto and alternative modes. First, higher proximity results in the presence of alternatives (to the auto) with higher speed and therefore lower travel time than walking and biking. At a certain high level of proximity, which is typically accompanied by a high level of neighborhood-based density, the high number of persons present—and therefore the high number of candidates for using public transit—lowers the expected subsidy per ride and causes government to be willing to invest in public transit, e.g. bus service. And at even higher densities, government is willing—for the same reason—to invest in lower travel times for public transit travel, i.e. reducing wait times by supplying greater bus frequency, and increasing speeds by supplying dedicated rights-of-way (e.g. BRT and rail). Concerning this impact of transit investment (and parking price, above), Salon et al. (*25*) note "Density is correlated with many of the other factors that we expect to affect VMT, including both land use factors and factors such as transit service and parking prices." Secondly, higher proximity affects the travel time difference between auto and all alternative modes. Proximity reduces the distance to destinations (as shown in detail below), thereby reducing the time to get to those destinations via any mode, thereby reducing the difference in travel time between auto and alternatives, and thus increasing the choosing of alternative modes by persons of certain incomes.

In summary of the above discussion of mode choice theory, proximity reduces VMT by reducing the choice of auto via increasing the cost of parking and reducing the difference in travel time between auto and alternative modes.

Proximity's Impact on Auto Travel Distance Secondly, given the choice of auto, proximity reduces the distance traveled to destinations of activity via a theory known as the intervening opportunities model. According to Schneider—who "developed the theory in the way it is presented today" (*14*)—the intervening opportunities model is based on the following assumptions (*41*):

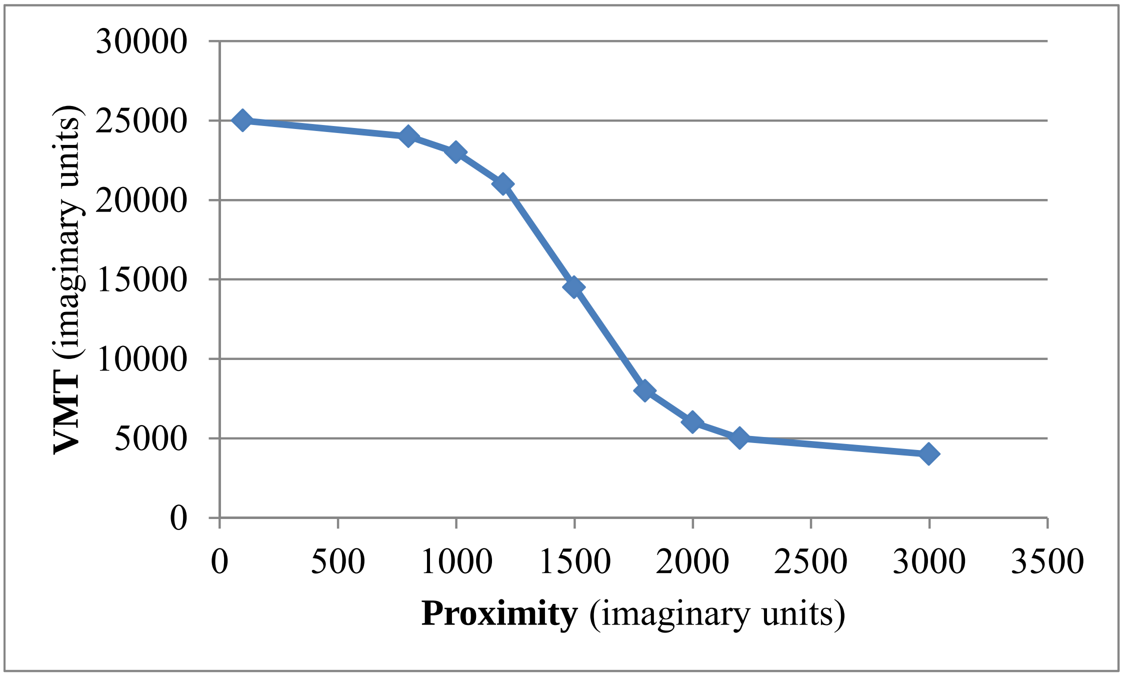
- "that the probability of a trip finding a terminal in any element of a region is proportional to the number of terminal opportunities contained in the element"
- "that a trip prefers to be as short as possible, lengthening only as it fails to find a [closer] terminal."

It follows then that the more destinations that lie within a given distance, the higher will be the probability of the traveler being satisfied by traveling that distance or less. Given that proximity is defined herein as "the degree to which the home is located near activity destinations", residents of homes with higher proximity will have shorter trips than those living in lower proximity areas. Therefore, given 1) the choice of auto (as established at the beginning of the paragraph), and 2) assuming a constant number of trips (assumption discussed in detail below), homes with higher proximity will have lower VMT.

Concerning the above assumption of a constant number of trips, the economic theory above of maximizing consumption leads one to conclude that drivers living in areas of high proximity will take advantage of being closer to destinations—each trip to an activity being less costly and requiring less time—by consuming more activities and (ignoring trip chaining) thus conducting more trips. It should be noted, however, that the trip-number-increasing effect of proximity being secondary to (i.e. in response to) the primary effect of proximity (shorter trip lengths), it is expected that proximity will decrease (overall) the VMT of auto users.

In summary, proximity reduces VMT 1) by reducing the choice of the auto mode (by increasing the cost of parking and reducing the difference in travel time between auto and alternative modes), and 2) by, given the choice of auto, reducing total trip distances.

The Expected Shape of VMT-Proximity Curve and Key Hypothesis Discovering the VMT impact at each level of proximity—the research objective of this dissertation—will result in VMT-proximity curves. It is expected that the "ideal" (using the Platonic meaning of the word) VMT-proximity curve flattens at both extremes of proximity, giving the curve an S-shape, i.e. somewhat similar to the shape of the Walls curve shown in Figure 3 above. At high levels of proximity, because there naturally exists a minimum household VMT (zero), it is expected that the curve will approach this minimum asymptotically—as is the case of other natural phenomena approaching a limiting value. Therefore, one expects a flattening VMT-proximity curve at the upper end of proximity. At low levels of proximity, because Schneider set the cumulative probability of trip-making equal to 1.0 (*41*), i.e. he assumed that the subject trip would be made regardless of how far one must travel to reach the first opportunity, the intervening opportunities model renders no maximum VMT. In reality, however, due to limited income and time discussed above, it is expected that persons living in remote areas will combine and forego trips. As a result, as one examines more-and-more remote (i.e. less-and-less proximate) households, a maximum VMT is expected. And, given this limiting value, it is expected that the curve will approach maximum VMT asymptotically, resulting in a flattening VMT-proximity curve at the lower end of proximity.

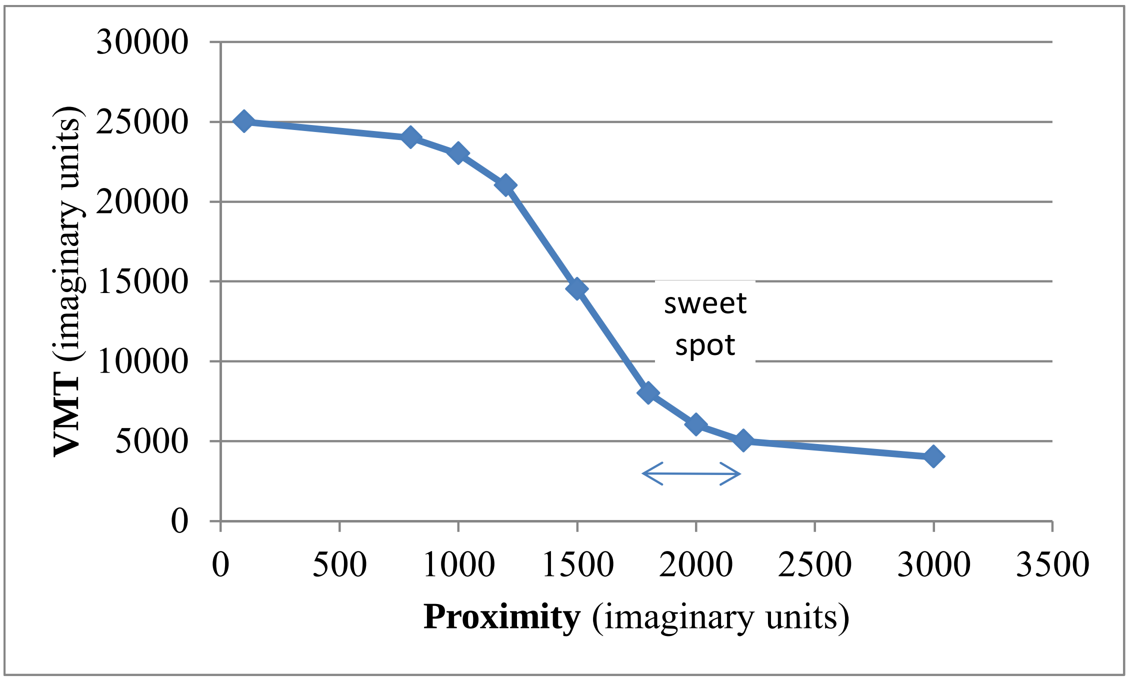


diagrams.xlsx

FIGURE 4 Expected Shape of VMT-Proximity Curve.

The expectation that the VMT-proximity curve is S-shaped—i.e. sloping sharply between the above two flat portions—is supported, in part, by the finding of Pushkarev and Zupan that there exists a bend in the transit-trips-vs.-residential-density curve at 7 dwelling units per acre (*7*). According to Pushkarev and Zupan, "…densities in the 2 to 7 dwellings per acre range produced only marginal use of public transportation…", whereas densities "of 7 to 30 dwellings per acre were necessary to sustain significant transit use..." This finding of a bend in the transit-use-vs.-density curve—below which density level transit use increases slowly with increasing density, and above which density level transit use increases more rapidly with increasing density—supports the existence of the first bend in the above-assumed S-shaped VMT-proximity curve—below which proximity level VMT decreases slowly with increasing proximity, and above which proximity level VMT decreases more rapidly with increasing proximity. And given the resulting steeply sloping section of the VMT-proximity curve, there must exist a second bend in order for the curve to flatten at higher proximity as theorized above.

The implication of the S-shape of the above theoretical VMT-proximity curve is that there exist points along the bend in the right-hand portion of the curve having *low* VMT and a *moderate* level of proximity as shown below. If, along the bend, the VMT is well below the average VMT and the proximity level is acceptable to a significant portion of the population (as discussed in the Impetus section above), then governments could promote housing development at this "sweet spot" proximity and thereby lower the average VMT of the population.



diagrams.xlsx

FIGURE 5 Expected Sweet Spot on VMT-Proximity Curve.

Therefore, *the key hypothesis of this dissertation is*:

> There exists a sweet spot on the VMT-proximity curve that has high VMT benefit *and* a proximity level acceptable to many households.

In order to test this hypothesis, a specific version of it must be developed. Based on the above S-curve, the sweet spot would be somewhat above the 50% proximity level. Assuming that 67% of the maximum proximity level is low enough to be acceptable to a significant portion of the population, and that 80% of the VMT benefit at maximum proximity is high enough to significantly lower the average VMT of the subject population, the following *specific key hypothesis will be tested*:

> The VMT benefit at 67% of maximum proximity is equal to or greater than 80% of the VMT benefit at maximum proximity.

The Expected Shape of VMT-Centrality Curve and Secondary Hypothesis It is expected for two reasons that the VMT-*centrality* curve is also S-shaped. First, centrality being a proxy for proximity, it is expected that the VMT-centrality curve will resemble the VMT-proximity curve and be, therefore, S-shaped as discussed above. Secondly, because centrality, unlike proximity, has a maximum value (i.e. being at center)—were one to move along an imaginary line bisecting the ideal metro (such line referred to in the literature as a "transect" (*35*)), moving away from the outer metro edge and approaching the center—one would expect the amount of proximity to level off before declining as one moves across (and therefore away from) the center. Therefore, VMT being related to proximity, it is expected that VMT would likewise level off—at some minimum average VMT value—as one approaches the metro center and, thus, the high-centrality-extreme of the VMT-centrality curve.

The implication of an S-shaped VMT-centrality curve is similar to the implication of an S-shaped VMT-proximity curve discussed above. The implication of the S-shape of the theoretical VMT-centrality curve is that there exist points along the bend in the right-hand portion of the curve having low VMT and a moderate level of centrality. If, along the bend, the VMT is below the average VMT and the centrality level is acceptable to a significant portion of the population, then governments could promote housing development at this "sweet spot" centrality and thereby lower the average VMT of the population.

Given that centrality is a proxy for proximity, a centrality-based hypothesis is secondary to the key proximity-based hypothesis. Based on the expected shape of the VMT-centrality curve, *the secondary hypothesis of this dissertation is*:

> There exists a sweet spot on the VMT-centrality curve that has high VMT benefit *and* a centrality level acceptable to many households.

In order to test this hypothesis, a specific version of it must be developed. Assuming that 67% of the maximum centrality level is low enough to be acceptable to a significant portion of the population, and that 80% of the VMT benefit at maximum centrality is high enough to significantly lower the average VMT of the subject population, the following *specific secondary hypothesis will be tested*:

> The VMT benefit at 67% of maximum centrality is equal to or greater than 80% of the VMT benefit at maximum centrality.

Measuring Proximity

Measuring Proximity in the Literature Researchers have measured proximity with various methods including methods based on centrality, density, thresholds, gravity model, etc. as discussed below.

Using Centrality as a Proxy for Proximity Given—as discussed above—that centrality is a proxy for proximity, some researchers have used centrality as an independent variable in their attempts to study the effect of land use on travel. In their 2010 study, Cao, Xu, and Fan (*30*) "classified households into four types of locations based on the network distance between households' residence and the city center point." They "divided the distance into four intervals: [0, 5] miles (called urban areas for simplicity), (5, 10] miles (called inner-ring suburbs), (10, 15] miles (called suburbs), and 15+ miles (called exurbs)." [The preceding quote is punctuated herein as originally written.]

Using Other Proxies of Proximity Salon et al. (*25*) analyzed the VMT impact of several proxies of proximity: land use mix, jobs-housing balance, and—considering the paths which join proximate locations—network connectivity.

Land Use Units Used in Proximity Measures When actually trying to measure proximity—as opposed to using the above proxies for proximity—authors usually represent the subject destinations (i.e. the locations to which the subject household is to some degree proximate) using discrete land use units, primarily population (i.e. the number of persons or households) and employment (i.e. the number of jobs). For example, for the simple proximity measurement discussed below—density—the discrete land use unit measured is usually population.

Some authors have used combinations of land use units in their proximity measures. Combining types of employment, Case (*4*) measured "Activity Location Units" (ALUs) within a certain threshold distance of the household, one ALU for each employee of a non-retail establishment and—to reflect the higher number of trips per employee attracted to retail businesses—three ALUs for each employee of a retail establishment. Zhou and Kockelman (*31*), on the other hand, combining population and employment, measured "person equivalents per acre", i.e. "zone population plus zone employment times the regional persons-per-job ratio…"

Using Density to Measure Proximity The most prevalent measure of proximity in the literature is density. Density is the number of discrete units of interest in a certain area—in this case, the area in which the subject household is located—divided by the size of that area (referred to herein as the "density area"). An example of density measurement is the number of households per square mile in the block group in which the subject house is located.

Many authors used density to measure proximity and explore its effects. Some used residential/population density:

- In 2000, Badoe and Miller reportedly examined the relationship between residential density and mode choice (*1*).
- In 2002, Holtzclaw et al. reportedly examined the relationship between residential density and VMT (*1*).
- In 2005, Golob and Brownstone (*21*) examined the relationship between population and housing unit density and VMT.

- In 2007, Ewing et al. reportedly examined the relationship between population density and VMT (*1*).

Some used employment density:

- In 1998, Boarnet and Sarmiento reportedly examined the relationship between retail employment density and nonwork auto trips (*1*).

- In 2005, Golob and Brownstone (*21*) examined the relationship between employment density and VMT.

The scope of density measurement, i.e. the size of the density area, varies by author. Many authors measured density on a neighborhood basis. In her dissertation, Shay (*24*) used "…a neighborhood typology…, which captures such environmental qualities as density, connectivity, and streetscape." Some authors, on the other hand, calculated density for entire metros. For example, in 2010, Cervero and Murakami examined the relationship between metro density and traffic (*6*).

Using Distance Thresholds to Measure Proximity Some authors measured proximity by examining the contents of the environment of a subject household, that environment measured out to a threshold distance. For example, Cervero and Duncan measured "the number of jobs in the same occupational category [as the human subject] within 4 miles of one's residence" *(3).* Yoon, Golob, and Goulias (*26*) measured employment within 10km and 50km of the subject census tract. This method will be referred to herein as "distance-threshold-based total opportunities."

Using the Gravity Model to Measure Proximity Measuring the environment of a home by using the gravity model is a fairly complex method in that all destinations in the modeled region are considered, not simply those within a certain threshold distance. Several authors have used the gravity model to measure the accessibility of homes. In 1993, Handy measured local and regional accessibility using gravity-based formulations (*27*). In 2001, Ewing and Cervero reportedly measured "destination accessibility" which was "represented by an accessibility index derived with a gravity model" (*1*). In 2003, Krizek (*32*) measured "regional accessibility" by entering retail employment into a gravity formula. The Access to Destinations Study (2010) used the gravity model to measure accessibility whereby "nearby things exert stronger attraction than those far away" (*8*). If applied with distances instead of travel times, these uses of the gravity model could be modified to measure proximity instead of accessibility.

Using Other Complex Methods of Measuring Proximity In addition to the gravity model, other complex methods of measuring proximity have been used. Khattak et al. (*5*) explored the relationship between various variables—including Claritas area types—and commute distance. For the 1995 Nationwide Personal Transportation Survey (NPTS) and the 2001 National Household Travel Survey (NHTS), Claritas measured proximity using a complex framework based on density (*9*). Every part of the nation was classified into five "Area Types." First, to reduce distortions caused by the delineation and size of census areas, Claritas divided the U.S. into 900,000 squares (called "grids"), each approximately 2 miles by 2 miles or 4 square miles in area. Secondly, each grid was assigned a "contextual density", the population density of the nine grids (3x3) for which the subject grid is the central grid, i.e. not the density of the subject grid itself. (Note that this is roughly equivalent to measuring the number of persons within a 3 mile threshold of the subject household.) Thirdly, based on that contextual density, each grid was placed into one of three density categories: low density, medium/low density, medium-to-high density. The first two of these categories were used as the first two (of the five) area types, which are therefore referred to by this author as the "Low Density Area Type" and "Medium/Low Density Area Type" (the Claritas labels for these two Area Types ["Rural" and "Town", respectively] being misleading). Fourthly, the "relational nature" of the grids in the third density category (medium-to-high) was identified, i.e. how a grid's contextual density compares to that of adjacent grids. Based on the relative contextual density of nearby grids, Claritas judged these grids as having one of two natures: a *central* nature (i.e. having the highest, or nearly the highest, contextual density in the vicinity), or a *surrounding* nature (i.e. having significantly lower contextual density than that of its "population center"). These surrounding grids comprise the third area type, which is therefore referred to by this author as the "Medium-to-High Density Surrounding Area Type" (imprecisely labeled "Suburban" by Claritas). Finally, Claritas split the grids with a central nature into the fourth and fifth area types, based on the contextual density of a grid's population center, placing those with medium density in the fourth area type, and placing those with high density in the fifth area type. Therefore, this author refers to these two area types as the "Medium Density Central Area Type" and the "High Density Central Area Type" (misleadingly labeled "Second City" and "Urban" respectively by Claritas).

In 2005, Bento et al. used annuli to measure proximity, another complex method. They computed "population centrality" (a property of the metro representing its spread, not to be confused with the "centrality," a property of the household examined in this dissertation) by "averaging the difference between the cumulative population in annulus *n* (expressed as a percentage of total population) and the cumulative distance-weighted population in annulus *n* (expressed as a percentage of total distance-weighted population)" (*18*).

Measuring Proximity in this Dissertation The above literature review discovered many different methods of measuring proximity:

- Centrality
- Other proxies (land use mix, jobs-housing balance, and network connectivity)
- Density
- Distance-threshold-based total opportunities
- Complex methods (gravity-based, Claritas area types, and annuli-based)

Although there is no perfect method of measuring proximity, several of these methods have unacceptable weaknesses. Given that even the most dense and varied neighborhood can satisfy only a small portion of the desired activities of mobile citizens, it is expected that neighborhood-based measures, such as network connectivity, have little relationship to VMT, an assumption supported by the literature as discussed below in section "The Empirical Impact of Proximity on VMT in the Literature." Concerning the gravity-based and annuli-based methods, their complexity prevents them from being readily understood by potential consumers of this dissertation, and the categorical nature of the Claritas method prevents it from being plotted, disqualifying these methods from being applied herein for creation of VMT-proximity curves.

On the other hand, centrality, i.e. the closeness to the center of the metro area, is a common and simple way of considering location. For example, someone might ask another person, "How far out do you live?", i.e. "How far do you live from downtown?" Likewise, the popularity and simplicity of density make it an attractive method for this dissertation. Density provides proximity over the density area. Although this density area is often smaller than the large regional area over which proximity is best measured (as discussed below), density over small areas is related to proximity over large areas, and is therefore useful for measuring that proximity. And finally, given the intuitive theory of intervening opportunities discussed above, distance-threshold-based total opportunities—which measures at least a portion of those intervening opportunities—is also an appropriate measure of proximity for this dissertation. Therefore, in this dissertation, VMT-proximity curves are developed using:

1) centrality
2) density, and
3) distance-threshold-based total opportunities.

The Empirical Impact of Proximity on VMT in the Literature

In this section, the findings in the transportation literature concerning the impact of proximity on VMT are summarized. VMT is a mathematical function of three determinants: 1) the mode of travel, 2) the distance of trips, and 3) the number of trips. The findings below are organized by the above three determinants of VMT plus a fourth category for VMT itself.

Relationship Between Proximity and Mode According to the literature, an increase in the usage of alternative modes does not necessarily indicate an equivalent decrease in the usage of autos. For example, according to Shay's review of the literature (*24*), "The increased proximity afforded by mixing residential, retail, and office land uses appears to support walking trips; however, it is less clear whether such trips complement or substitute for existing trips that rely on motorized modes (Ewing and Cervero, 2001; Handy, 2006)." And Salon et al. (*25*) state that "We expect that as transit ridership increases, VMT will decrease, but the effect is likely to be less than one-to-one, both because new transit trips do not always replace car trips and because of latent demand for road space…."

Accordingly, some of the literature points to the small impact of alternative modes on VMT. Concerning the effect of walking on VMT, according to Salon et al. (*25*):

> "There have been a handful of studies that identify the VMT effect of walking, and the results have been mixed. In a study of Portland, Oregon, Parsons Brinkerhoff (1993) found an elasticity of VMT with respect to a measure of pedestrian quality of -0.19."
>
> "Kitamura, et al. (1997) found that the presence of sidewalks in the neighborhood was associated with a 0.14 percent decrease in vehicle trips."

Concerning the effect of public transit on VMT, according to Salon et al. (*25*):

> "For fare, frequency, and service miles/hours, the literature provides evidence on the relationship of these characteristics of a transit system to transit ridership, but the effect on VMT is not quantified."
>
> "Paulley et al. (2006) is one of the few studies that examined links from service characteristics to car use, and they found that the elasticity of automobile mode share with respect to bus transit fare was about -0.05, approximately one-tenth the fare elasticity estimate of transit ridership."

Concerning the effect of biking on VMT, according to Salon et al. (*25*), "To our knowledge, the link between increased bicycling and VMT reduction has not been empirically quantified."

Relationship Between Proximity and Trip Distance Using 1995 NPTS data, Khattak et al. (*5*) explored the relationship between various variables—including Claritas area types—and commute distance. As discussed above, the NPTS defined five area types: 1) Low Density (a.k.a. "Rural"); 2) Medium/Low Density (a.k.a. "Town"); 3) Medium-to-High Density Surrounding Area (a.k.a. "Suburban"); 4) Medium Density Central Area (a.k.a. "Second City"); and 5) High Density Central Area (a.k.a. "Urban"). Based on the detailed description of their composition in the "Measures of Proximity" section above, it appears that these area types are ordered according to proximity—from "Low Density" areas (representing the lowest proximity), to "High Density Central" areas (representing the highest proximity). Based on this apparent order of proximity and the VMT theory of this dissertation, one would expect these area types to be also ordered according to commute distance—from "Low Density" areas having the longest commute distances, to "High Density Central" areas having the shortest commute distances. Although this order held when simply examining average distances by area type, this order did not hold when controlling for other variables, e.g. controlling for income to account for any "spatial mismatch hypothesis" effects. Surprisingly, the coefficients from Khattak's weighted least squares regression model indicated that workers living in Medium-to-High Density Surrounding Areas (the base area type) had *shorter* commute distances than those living in High Density Central areas (coefficient +1.91), ceteris paribus.

Relationship Between Proximity and Number of Trips Examining metro density, Cervero and Murakami (*6*) found "traffic-inducing effects of denser urban settings having denser road networks and better local-retail accessibility (indirect effect elasticity =0.223…)." It is assumed that the higher VMT which they found is due, in part or whole, to increased trip making. More directly, Shay (*24*) found a positive relationship, ceteris paribus, between number of trips and both 1) living in and around the CBD, and 2) residential density.

Relationship Between Proximity and VMT Just as the literature records the relationship between proximity and the three components of VMT—1) mode, 2) trip distance, and 3) number of trips—it also records the relationship between proximity and VMT itself. Some studies examine the effect of density on VMT:

- In 1997, Ross and Dunning (*23*) calculated VMT by various levels of population density but they did not control for any other variables.
- In 2002, Holtzclaw et al. reportedly found that lower residential density is the cause of higher auto ownership and vehicle miles traveled (VMT) (*1*).
- In 2005, Golob and Brownstone found: "Comparing two households that are similar in all respects except residential density, a lower density of 1,000 housing units per square mile [block group measure] implies a positive difference of almost 1,200 miles per year…" (*21*)

- In 2007, Ewing et al. reportedly found "a 0.152 percent reduction in VMT from a 1 percent increase in population density on the basis of their longitudinal model…" (*1*)

Some studies examine the effect of more complex measures of proximity on VMT. In 2005, Bento et al. found that "population centrality" (a measure of metro "spreadness" differing from the household "centrality" of this dissertation, as described above), by itself, had a modest effect on VMT:

> "Population centrality, which affects average VMTs only through its effect on vehicle choice, has a slightly larger, but still modest, effect. A 1% increase in population centrality reduces average annual miles driven by 1.5% when New York is removed from the sample. As we report elsewhere (Bento et al., 2003), the 10% increase in population centrality in the sample without New York reduces annual average VMTs by approximately 300 miles per year—approximately half the size (in absolute value) of a 10% increase in household income." (*18*)

Some studies record the relationship between distance-threshold-based total opportunities and VMT. According to Cervero and Duncan in their article "Which Reduces Vehicle Travel More: Jobs-Housing Balance or Retail-Housing Mixing?", job proximity ("the number of jobs in the same occupational category [as that of the human subject] within 4 miles of one's residence") is a powerful VMT reducer. They also found that retail and service proximity is a VMT reducer, but to a lesser extent than job proximity (*3*).

Scope of Measurement of Proximity Regardless of which of the above methods is used to measure proximity, the scope of that measure must be established. According to Krizek (*32*), "It is…important to distinguish between the effects of urban form at the neighborhood scale versus the regional scale." In her writing, Handy differentiates between these two (*27*):

- "The amount that a person travels is influenced by both the character of the particular community in which he or she lives and the spatial structure of the region of which that community is a part."

- "Neotraditional developments with high levels of local accessibility…will have a greater effect on nonwork travel when located at the edges of the region than when they are located within the region surrounded by highly developed areas."

Although a) proximity measures with a neighborhood scope, and b) proximity measures with a regional scope are both related to VMT, the literature indicates that regionally-based proximity has a greater impact on VMT than does neighborhood-based proximity:

- According to Badoe and Miller, Ewing found in his 1995 analysis that "good regional accessibility was found to cut down on household vehicular travel to a far greater extent than did localized density of mixed use" (*2*).

- Measuring density on the metro level, Cervero and Murakami (*6*) found "population densities are shown to be strongly…associated with VMT per capita (direct effect elasticity =-0.604)…"

- According to Shay (*24*), "Ewing (1995) determined total travel to be a function of regional access, and thus largely beyond the power of individual neighborhoods to shape."

- Both Badoe and Miller (2000) and Ewing and Cervero (2001) reportedly note "the futility of increasing density in the middle of nowhere as a policy to reduce VMT" (*1*).

- As reported in TRB Special Report 298:

    - Concerning the immediate environment of a household, in 2002 Bagley and Mokhtarian found "little remaining effect of neighborhood type on VMT after controlling for attitudes, lifestyle preferences, and sociodemographic variables" (*1*).

    - In 2008, Arrington and Cervero concluded "that the location of a TOD in a region—its accessibility to desired locations—and the quality of connecting transit service are more important in influencing travel patterns than are the characteristics of the TOD itself (e.g., mixed uses, walkability)" (*1*).

- According to Cervero and Duncan (*3*):

  - "While we measured job accessibility indices within radii of 1 to 9 miles around survey respondents' residences, the best-fitting estimates were for 4-mile radii;"
  - "As with the study of job accessibility, the 4-mile radius provided the best statistical fit for estimating the influences of retail-service accessibility levels on the VMT of tours for shopping and personal services."

- Although not large enough to provide a truly regional scope, census tracts provide a scope of analysis larger than the neighborhood. According to Yoon, Golob, and Goulias (*26*):

  - "…household density measured in census tracts explained better the indicators used here [non-motorized travel, high-occupancy-vehicle usage, and solo driving] than household density measured using block groups."

Scope of Measurement of Proximity in this Dissertation Given that the literature indicates that large-scope proximity (i.e. regionally-based, census-tract-based) has a greater impact on VMT than does neighborhood-based proximity, in this dissertation proximity will primarily be measured at a large scope. Secondarily, neighborhood-based proximity will be measured to determine and control for its impact.

Summary of Empirical Impact of Proximity on VMT in the Literature In summary, as one might expect, the existing literature generally indicates that proximity tends to 1) increase the usage of alternative modes, 2) decrease trip length, and 3) increase the number of trips, overall lowering VMT. In addition, as discussed above, the literature contains a few findings that perhaps controvert conventional wisdom:

- Workers living in Medium-to-High Density Surrounding Areas had *shorter* commute distances than those living in High Density Central Areas, ceteris paribus.

- Large-scope proximity (i.e. regionally-based, census-tract-based) has a greater impact on VMT than does neighborhood-based proximity.

Socio-economics

As briefed in the overview above, it is theorized that work status, income, gender, age, number of persons in the household, and disabilities are determinants of annual household VMT. The term “socio-economics”, meaning the “combination of social and economic factors” (*40*), covers these determinants.

Socio-economics in the Literature

Many studies which examined the impact of the built environment on travel and VMT used socio-economics as control variables in an attempt to isolate the impact of the built environment:

- Most of the studies reviewed in Special Report 298 (*1*) treated socio-economics as control variables. Holtzclaw et al. (2002), Bagley and Mokhtarian (2002), Frank et al. (2007), Bhat and Guo (2007), Brownstone (2008), and Brownstone and Golob (2009) all controlled for socio-economics in their analyses. Holtzclaw used household size and income as control variables, whereas Bhat and Guo used various household characteristics, and found that household income strongly affects car ownership (and therefore travel).

- In their analysis of transportation-land use literature, Badoe and Miller (*2*)—referencing Peat Marwick & Mitchell (1975); Schimek (1995); Loutzenheiser (1997)—stated that “…socioeconomic factors…such as income, age, gender, occupation, etc., have a significant impact on travel behavior.”

- Cervero and Duncan (*3*) controlled for socio-economics when they examined the relative impact of job proximity and retail/service proximity on travel. Their control variables reflected income level, type of employment, age, ethnicity, motor vehicle ownership, driver licensing, student status, and gender.

Two of the studies reviewed in Special Report 298 (*1*)—Handy (2005) and Ewing and Cervero (2001)—examined the relative importance of socio-economics and the built environment in impacting travel. Ewing and Cervero found that “socioeconomic factors are dominant in trip frequency decisions, whereas the built environment appears to be more influential with respect to trip length...”

Socio-economics in this Dissertation

Given the theoretical and documented impact of socio-economics on travel and VMT, control variables reflecting work status, income, gender, and age, will be included in the models developed for this dissertation.

Auto Ownership

Household VMT being a mathematical product of the number of vehicles in the home and the amount of travel per vehicle, socio-economics and proximity impact VMT, in part, through the step of auto ownership. Without auto ownership, there is no VMT. Given the time advantage and high cost of auto travel, as discussed above, higher incomes increase the tendency of a household to travel by auto, including the necessary and expensive step of purchasing and insuring an auto. And given that proximity (with its above-described companions: transit service level and parking costs) reduces the advantage of auto travel by reducing the travel time difference between auto and alternative modes (as described above), greater proximity is expected to decrease the tendency of a household to purchase and insure an auto.

Auto Ownership in the Literature

The literature discusses the role of auto ownership in the relationship between urban design and VMT. According to Badoe and Miller (*2*):

> "…auto ownership is a critical "intermediate link" between household location choices (where to live, where to work) and their subsequent activity/travel decisions."
>
> "Thus…a proper specification of the urban form - travel demand interaction - requires including auto ownership as an endogenous component of the system."

Likewise, according to Schimek (*13*):

> "…given a neighborhood of a certain density, households choose the number of vehicles to hold (own, lease, etc.) and then decide the number of motor vehicle trips or total vehicle travel distance…."

Bento et al., in their 2005 study of the effects of urban spatial structure on travel demand, estimate a VMT model in two parts (*18*):

> "The first part is a multinomial logit model that explains whether the household owns zero, one, two, or three or more vehicles. We then study the determinants of annual VMTs per vehicle separately for households that own one, two, or three or more vehicles."

Likewise in Shay's dissertation examining the travel impact of the various types of urban environments in the Charlotte metro (*24*), "Path analysis is used to examine the relationship of environment with travel—both directly, and indirectly through auto ownership."

Auto Ownership in this Dissertation

Given that auto ownership is an effect of proximity, modeling VMT without mixing causes and effects in the set of independent variables, may be achieved in two ways: 1) via two-part modeling (as done by Bento and Shay above), or 2) by *excluding* auto ownership from a set of independent variables that includes socio-economics and proximity. If one desires to examine the components of the impact of socio-economics and proximity on VMT—i.e. what portion of that impact is exercised through auto ownership, and what portion of that impact is exercised through mileage-per-auto—then one would use two-part modeling to do so. If, however, one is simply interested in the relationship between proximity and VMT—as in this dissertation—one can account for the endogeneity of auto ownership by simply excluding auto ownership from the set of independent variables in a one-part model, including instead only exogenous variables such as proximity and socio-economics. Consequently, for the sake of simplicity, one-part models that exclude “number of vehicles” as an independent variable (IV) will be employed in this dissertation.

Note that—due to the strong logical and empirical relationship between the presence of vehicles and the presence of drivers in a household—the models in this dissertation will also exclude “number of drivers” as an independent variable.

Time-Based Accessibility

As briefly discussed in the overview above, it is expected that access to high-speed roadways tends to increase the VMT of households. Comparing two households with their only difference being the speed of the roadways in their environments (i.e. they have the same [distance-based] proximity to destinations, but one is served by high-speed roadways, the other is served by low-speed roadways), one expects the household with high-speed highways to take advantage of those highways and sometimes choose distant but desirable destinations because they can be reached quickly. Therefore, one expects households in an environment of high-speed highways to have higher VMT than households in other environments.

Given that proximity, which reflects distance-based access to destinations, will be included in the set of independent variables in this dissertation’s models, the *extra* access to distant destinations provided by high-speed roadways can be represented in these same models by including a variable measuring the destinations within a certain travel *time* of the subject household, i.e. “time-based accessibility”, e.g. “number of persons within 20 minutes.” Whereas placing accessibility, without proximity, in a model would be expected to result in accessibility being negatively related to VMT (because of the general similarity between accessibility and proximity), pairing accessibility with proximity in a model is expected to result in accessibility being positively related to VMT and thereby reflecting the impact of high-speed roadways.

Travel Mode Biases ("Self-Selection")

As briefed in the overview above, it is assumed that, to a certain degree, people have individual biases toward the various travel modes, and therefore choose where they live, in part, in accordance with those biases (known in the literature as "self-selection"). Special Report 298 relates that "Boarnet and Crane (2001), among others, note that the observed correlation between higher-density neighborhoods and less automobile travel may be due in part to the fact that some residents who dislike driving and prefer transit or walking or bicycling may have self-selected into neighborhoods where these travel options are available" (*1*). For example, people who do not like to drive will tend to choose to live in intensely urban places such as Manhattan. Therefore, were one to attribute the walking habits of New Yorkers solely to the proximity of their homes to destinations (and to the associated alternative transportation infrastructure), one would be overestimating the impact of that proximity.

Travel Mode Biases / Self-Selection in the Literature

Researchers have found that self-selection can significantly distort the results of research. According to Cao et al. (*30*):

> "Using a sample of 1,903 households in a travel survey in Austin, TX, Zhou and Kockelman (Zhou and Kockelman, 2008) applied Heckman's sample selection model. After controlling for self-selection, they found that households in suburban areas were likely to drive 27% more per day than those in urban areas. They

> concluded that self-selection explained 42% of the total influence of the built environment on vehicle miles traveled (VMT)."
>
> "Using a 2003 data of respondents living in four traditional and four suburban neighborhoods in Northern California, Cao (Cao, 2009)…employed a sample selection model to quantify the influence of neighborhood type itself on driving behavior. He concluded that about 24% of the total influence of neighborhood type on vehicle miles driven resulted from residential self-selection."

According to Zhou and Kockelman (*31*):

> "Depending on model specification used, results suggest that at least half (58% to 90%) of the differences in vehicle miles traveled observed between similar households living in CBD or urban versus rural or suburban neighborhoods of Austin is due to the location or treatment itself, whereas self-selection of such treatment (by households that wish to meet special travel needs or preferences) accounts for the remainder."

In response to this threat of self-selection to the validity of analyses, authors have adopted various analytical methods to address this issue, as follows.

Walls, Harrington, and Krupnick (*20*) used a "selectivity correction term" to address self-selection: "Using techniques developed by Heckman (1978, 1979), Dubin and McFadden (1984) developed a selectivity correction term for use in situations when the discrete choice probabilities are logit and the errors in the regression equations (the VMT equations here) are normally distributed."

Golob and Brownstone (*21*) used simultaneous equations to deal with self-selection:

> "We adopt a more direct approach to the problem of selectivity bias in disaggregate studies. The approach is to apply a simultaneous equations model in which residential density, vehicle usage, and fuel consumption are joint endogenous variables. In this way we can model socioeconomic and demographic effects on each of these three endogenous variables, while simultaneously capturing the direct effects of residential location on the vehicle usage and energy consumption."

In a later work, however, Brownstone (*12*) indicated that self-selection could be dealt with via the use of extensive independent variables in a traditional regression model:

> "Many studies with disaggregate data attempt to control for observable differences between people living in high and low density areas using regression methods. These studies are only valid to the extent that these people differ only on observable characteristics. Therefore studies like Bento et. al. (2005) which includes a rich set of household socioeconomic characteristics should be less affected by self-selection bias."

Travel Mode Biases in this Dissertation

As in the case of Bento et. al. above, travel mode biases will be addressed in this dissertation by including several key socio-economic variables in the VMT models. To the degree that any unobserved travel biases are correlated with these socio-economic variables, the effect of these biases will be captured in the coefficients of the subject socio-economic variables, and will consequently distort less the coefficients of the proximity variables.

Subsets of VMT: Type of Travel

To discover the proximity-travel relationship, some researchers analyze travel by dividing it into various types of trips. Handy isolated shopping travel (*27*):

- She found a relationship between two gravity-based measures and trip distance in her 1993 analysis—"In both cases [regional and local], shopping distance decreases with increasing accessibility."
- She found "The relationship between regional accessibility and shopping trips per person was virtually nonexistent…, as was the relationship between local accessibility and shopping trips per person…"

According to Badoe and Miller (*2*), Ewing—in investigating "the effects of land-use patterns on household travel behavior"—classified tours as either work related or non-work related. Likewise, Salon et al. (*25*) surmised the importance of commute trips in explaining VMT: "The effect of employment accessibility on VMT appears to be related to the large contribution of longer trips (presumably from commuting) to VMT." On the other hand, according to Krizek (*32*), "Hanson (1980) stressed the importance of analyzing work and nonwork travel jointly, because separating trips by type fails to capture the touring travel behavior that we know exists."

In this dissertation, because energy independence is affected by total VMT (i.e. the sum of all subsets of VMT), household VMT will be examined as a whole, not broken down into subsets such as shopping or work. Such detail is not necessary for the stated research objective of the dissertation: to discover the VMT impact of each level of proximity.

CHAPTER IV

EMPIRICAL ANALYSIS

Original empirical research was conducted to meet the research objective of this dissertation—to discover the VMT impact of each level of proximity—and fulfill the purpose of this dissertation identified above:

> to discover the VMT impact of each level of proximity in order to help government identify key locations for housing development, and thereby lower VMT and reduce dependence on foreign oil.

Toward that end, three research efforts were completed that examine the impact of proximity on VMT at each proximity level:

1. In the first effort, the proximity-VMT relationship was examined using data from across the nation and density as the measure of proximity.

2. In the second effort, designed to parse the findings of the first effort, the national data was used to explore the relationship between density and the usage of alternative modes.

3. In the third effort, the proximity-VMT relationship was examined using data from one region—Hampton Roads, Virginia—and distance-threshold-based total opportunities and centrality as measures of proximity.

Effort #1: Identifying Key Proximity Levels for VMT Using National Data

In order to identify key locations for development, Effort #1 was designed to discover VMT impact by proximity level. The analysis was conducted using a national dataset and density as the measure of proximity.

Data Preparation

All data for this effort came from the 2009 National Household Travel Survey (NHTS), using the special "DOT" file which contains additional variables not available from the NHTS website. Given the immense size of the data set—150,147 households, 308,901 persons, and 309,163 vehicles—I randomly extracted a subset of data that would be both manageable by modern PCs and render statistically significant results. In order to retain at least 100 records for the least prevalent density level, I randomly selected 9,961 household records.

Choice of VMT Variable Annual VMT was chosen over daily VMT as the dependent variable for three reasons. First, the fact that annual VMT is more familiar to people than daily VMT renders annual VMT research more easily applied. For example, people have a better idea of the significance of 15,000 miles per year than 15 miles per day. Secondly, annual VMT is more stable than daily VMT, the latter being subject to weather, temporary illness, holidays, etc. And finally, annual VMT is more suitable for easily-interpreted ordinary least squares (OLS) regression models.

Given that the NHTS contains three different types of estimates of annual VMT—1) odometer-based, 2) self-reported, and 3) estimate based on sample day—a particular type had to be chosen for this effort. Concerning VMT based on the single odometer reading on the NHTS, the NHTS literature reads:

> "Unfortunately, not all vehicles had an odometer reading recorded. Furthermore, of those that had their odometer reading recorded, the quality of some of the odometer readings is less than desirable." (*43*)

Given the difficulty of accurately converting the odometer readings included in the NHTS into annual VMT, this type of mileage data was rejected. The annual VMT estimate ("BESTMILE") calculated by the Oak Ridge National Laboratory (ORNL) for the NHTS using various sources including miles driven on sample day was estimated using NHTS socio-economic variables. Given this, using BESTMILE as my dependent variable would have inappropriately placed that socio-economic information on both sides (dependent and independent) of my regression. Concerning the remaining type, self-reported annual VMT by vehicle ("ANNMILES"), excerpts from a comparison to FHWA's *Highway Statistics* performed for the NHTS follows (*43*).

TABLE 3 Comparison of ANNMILES to FHWA's *Highway Statistics*

Source and Item	**Average Miles per Vehicle**
Highway Statistics (2008)	
Passenger Cars	11,788
Other 2-Axle, 4-Tire Vehicles	10,951
All	11,432
NHTS "ANNMILES" (2008-2009)	
Automobile/car/station wagon	10,054
Van (mini, cargo, passenger)	11,030
Sports utility vehicle	11,584
Pickup truck	9,891
All	10,088

Tables.xlsx

Based on the elimination of the first two types of annual VMT estimates, and the similarity between ANNMILES and the *Highway Statistics*, self-reported annual VMT by vehicle (ANNMILES) was used, aggregated to the household level.

Handling Missing Data Deleting those household records for households with vehicles having missing ANNMILES values (1,002 households) created the dataset of 8,959 households (9,961 – 1,002 = 8,959) used in the analysis. Concerning the independent variables of policy interest—the density variables—only one of the 8,959 household records had missing density data. Median densities of population and employment were assumed for that record. Concerning the set of control variables, missing household income was treated as a category of income, as shown in the "Descriptive Statistics" table below.

Data Validity Given that the NHTS (and its predecessor the National Personal Transportation Survey (NPTS)) has been conducted several times (1969, 1977, 1983, 1990, 1995, 2001, and 2009), is based on telephone interviews, and is financed by the federal government, the 2009 NHTS data tends to be valid. The annual household VMT used as the dependent variable in this analysis is based on the respondents' estimate of annual miles for each household vehicle. Although most people do not know exactly how many miles their vehicles have been driven during the past 12 months, it is expected that the error in those estimates is random and not correlated with any of the independent variables in the analysis. The key independent variables measuring proximity discussed below (population and employment density by census tract), having been prepared by Nielsen Claritas, are assumed to be reliable.

Although the usage of a robust set of independent variables in this effort's models removes any requirement that the subject sample dataset reflect exactly the population data, the following table demonstrates the similarity between the weighted full NHTS dataset and the unweighted analysis dataset.

TABLE 4 Similarity Between Full Dataset and Analysis Dataset

		Full Dataset (150,147 HHs)		Analysis Dataset (8,959 HHs)	
Household Variable	**NHTS Name**	**Unweighted Mean**	**Weighted Mean**	**Unweighted Mean**	**Weighted Mean**
Driver Count	DRVRCNT	1.80	1.73	1.77	1.73
Person Count	HHSIZE	2.34	2.47	2.29	2.36
Vehicle Count	HHVEHCNT	2.05	1.86	2.00	1.81
Unit Owned	HOMEOWN	87%	67%	87%	67%
Adult Count	NUMADLT	1.89	1.88	1.86	1.83
Worker Count	WRKCOUNT	0.93	1.09	0.91	1.06

Tables.xlsx

Selection and Preparation of Independent Variables Independent variables (IVs) were chosen for this effort's regression based on the theory and literature discussion in the "Preparation" section above, as summarized in the following table. The selection of an IV for each determinant is discussed below.

TABLE 5 Summary of Theorized Determinants of Annual Household VMT

Determinant	Universe
Proximity	Household
Internet Connectivity	Person, Household
Time-Based Accessibility	Household
Public Transit Service Level	Household
Travel Mode Biases ("self-selection")	Person
Socio-economics	
Work Status	Person
Income	Person, Household
Gender	Person
Age	Person
Number of Persons	Household
Disabilities	Person

Tables.xlsx

Proximity As discussed in the "Measuring Proximity" section above, 1) centrality, 2) density, and 3) distance-threshold-based total opportunities are desirable methods of measuring proximity due to their ease-of-interpretation and theoretical relationship to VMT. Given that—of these three—only density is readily available in the NHTS, this national analysis was performed using density. (Although distance-threshold-based total opportunities can be measured using additional efforts of reasonable difficulty for one metro area—as shown in this dissertation's third effort below—this is too difficult in a nation-wide analysis such as this first effort.) Given that—as described in section "The Empirical Impact of Proximity on VMT" above—1) the literature indicates that regionally-based proximity has a greater impact on VMT than does neighborhood-based proximity, and 2) Yoon, Golob, and Goulias found that "household density measured in census tracts explained better [non-motorized travel, high-occupancy-vehicle usage, and solo driving] than household density measured using block groups" (*26*), the NHTS variables based on census tracts (HTEEMPDN for employment density and HTPPOPDN for population density) were chosen over that of block groups (HBPPOPDN) to prepare the two sets of density IVs, one for employment and one for population. Because the NHTS variables contain values indicating ranges (e.g. in HTEEMPDN, "75" represents the density range 50-99 employed persons per square mile), a binary variable (e.g. "50-99 Employed Persons /sqmi, tract") was prepared for each range. Because the dataset includes a set of variables based on employment locations—the destination of most trips—it is richer than the typical transportation dataset containing only population densities.

In addition—given that the larger the metro area, the greater the distances to potential destinations—the NHTS variable MSASIZE was used to prepare the IVs "In MSA <1m Persons" and "In MSA/CMSA > 1m Persons", with basis variable "Not in MSA or CMSA." ("Basis" variables are those variables [from a set of binary variables covering the whole dataset] excluded from regressions to avoid over-specifying models. The impact of included variables from the set is measured as compared to the excluded basis variable.) Finally—given the popularity of land use mix analyses in the literature—a variable was prepared to reflect any interaction between high population density and high employment density. Of the eight population density levels, the top three levels (i.e. 4,000+ persons per square mile) were considered to be "high" population density. Likewise, of the eight employment density levels, the top three levels (i.e. 1,000+ employed persons per square mile) were considered to be "high" employment density. Therefore, the land use mix interaction variable "Pop Density >4k and Emp Density >1k" was prepared to identify those households which lie in census tracts which have *both* of these high density levels.

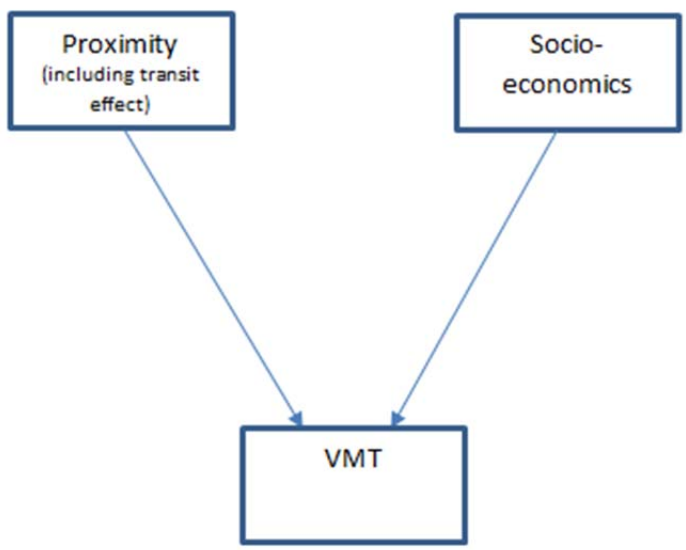
Internet Connectivity Concerning "internet connectivity" in the above table of determinants, the NHTS variable WEBUSE was used to calculate "Persons 16+ Used Internet Almost Every Day" and "Persons 16+ Never Used Internet in Past Mo." Likewise, the NHTS variable PURCHASE was used to calculate "Internet Purchases in Past Month."

Time-based Accessibility, Public Transit Service Level, and Travel Mode Biases Concerning time-based accessibility, public transit service level, and travel mode biases, no NHTS variables were available to directly measure these determinants. Concerning transit service level, however—as discussed in the "VMT Theory" section above—density is highly related to transit service. Therefore, the impact of transit service on VMT is part of the impact of this effort's density variables, and is measured therefore—along with the other impacts of density—in the coefficients of the density variables. Concerning travel mode biases, these biases (or "self-selection") were addressed in this effort in the Brownstone (*12*) manner discussed in the Preparation section above, i.e. by including several key socio-economic variables in the model.

Socio-economics Concerning socio-economics, work status, gender, age, and number of persons were collectively represented by using the NHTS variables R_AGE, R_SEX, and WORKER to prepare the IVs "Male Workers (Age 16+)", "Female Workers (Age 16+)", "Male Non-Workers (Age 16+)", "Female Non-Workers (Age 16+)", and "Persons Age 5 thru 15." (The NHTS does not record the age of household members younger than 5.)

Household income was represented 1) by using the NHTS variable HHFAMINC to prepare the set of binary income IVs (“HHFAMINC $20,000-$39,999”, “HHFAMINC $40,000-$59,999”, etc.), and 2) by using the NHTS variable HOMEOWN to prepare the binary variable “Home Owned.” Disabilities were represented by using the NHTS variable MEDCOND to prepare the IV “Persons 16+ Having MEDCOND.”

A drawing of the relationship between the dependent variable and key independent variables is shown below.



key relationships1.png

FIGURE 6 Key Relationships- Effort #1

TABLE 6 Descriptive Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
ANNMILES (annual household VMT)	8,959	19,011	18,381	0	265,200
Derived Total Household Income					
basis: HHFAMINC <$20k	8,959	0.15	0.36	0	1
HHFAMINC missing	8,959	0.07	0.26	0	1
HHFAMINC $20,000-$39,999	8,959	0.21	0.41	0	1
HHFAMINC $40,000-$59,999	8,959	0.17	0.38	0	1
HHFAMINC $60,000-$99,999	8,959	0.21	0.41	0	1
HHFAMINC $100,000+	8,959	0.18	0.39	0	1
		1.00			
Home Owned	8,959	0.87	0.33	0	1
All Household Members (Age 5+)					
Male Workers (Age 16+)	8,959	0.47	0.57	0	3
Female Workers (Age 16+)	8,959	0.44	0.55	0	3
Male Non-Workers (Age 16+)	8,959	0.35	0.50	0	4
Female Non-Workers (Age 16+)	8,959	0.54	0.55	0	3
Persons Age 5 thru 15	8,959	0.26	0.66	0	5
Persons 16+ Having MEDCOND	8,959	0.22	0.46	0	3
Internet Purchases in Past Month	8,959	2.48	5.55	0	200
Persons 16+ Used Internet Almost Every Day	8,959	1.01	0.92	0	5
Persons 16+ Never Used Internet in Past Mo.	8,959	0.46	0.68	0	5
Size of Area of Residence					
basis: Not in MSA or CMSA	8,959	0.21	0.41	0	1
In MSA <1m Persons	8,959	0.31	0.46	0	1
In MSA/CMSA >1m Persons	8,959	0.48	0.50	0	1
		1.00			
Population Density of HH Census Tract					
basis: <100 Persons/sqmi, tract	8,959	0.16	0.37	0	1
100-499 Persons/sqmi, tract	8,959	0.19	0.39	0	1
500-999 Persons/sqmi, tract	8,959	0.10	0.30	0	1
1,000-1,999 Persons/sqmi, tract	8,959	0.14	0.34	0	1
2,000-3,999 Persons/sqmi, tract	8,959	0.19	0.39	0	1
4,000-9,999 Persons/sqmi, tract	8,959	0.18	0.39	0	1
10,000-24,999 Persons/sqmi, tract	8,959	0.03	0.18	0	1
25,000+ Persons/sqmi, tract	8,959	0.01	0.12	0	1
		1.00			
Employment Density of HH Census Tract, by Place of Employment					
basis: <50 Employed Persons /sqmi, tract	8,959	0.24	0.42	0	1
50-99 Employed Persons /sqmi, tract	8,959	0.07	0.25	0	1
100-249 Employed Persons /sqmi, tract	8,959	0.12	0.33	0	1
250-499 Employed Persons /sqmi, tract	8,959	0.13	0.33	0	1
500-999 Employed Persons /sqmi, tract	8,959	0.15	0.36	0	1
1,000-1,999 Employed Persons /sqmi, tract	8,959	0.14	0.35	0	1
2,000-3,999 Employed Persons /sqmi, tract	8,959	0.09	0.29	0	1
4,000+ Employed Persons /sqmi, tract	8,959	0.06	0.24	0	1
		1.00			
Pop Density >4k and Emp Density >1k	8,959	0.16	0.37	0	1

Tables.xlsx

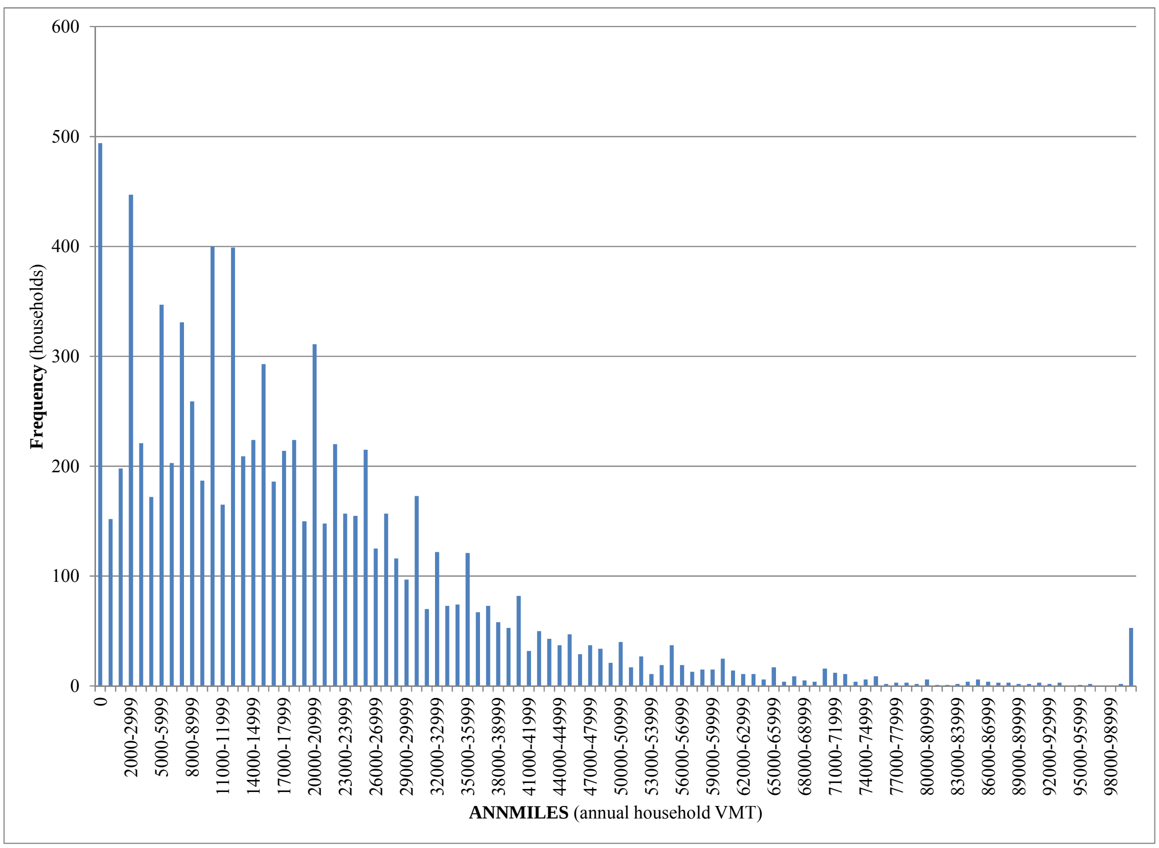
Descriptive Statistics

The descriptive statistics in the above table provide a detailed view of American households. Given the difference between weighted and unweighted values in the "Similarity" table above, some of the statistics in the above table of unweighted values will differ from actual average national values. Concerning the dependent variable, the average household VMT is approximately 19,000 miles.

Concerning the independent variables, the extensive presence of binary variables in the dataset allow for easy categorization of the dataset's households. With approximately 2/5ths of the households having lower income and an equal share having higher income, median household income is approximately $50,000. (Fortunately, only 7% of the household records are missing income information.) Although home ownership in the dataset is very high (87%), note that the weighted value shown in the "Similarity" table above is significantly lower (67%). Half the households are located in MSA/CMSAs with more than 1 million population, the other half in less populous areas. With approximately 2/5ths of the households being located in lower population-density tracts and an equal share located in higher density tracts, the median household census tract population density is in the 1,000-2,000 persons per square mile range. Likewise, with approximately 2/5ths of the households being located in lower employment-density tracts and an equal share located in higher density tracts, the median household census tract employment density is in the 250-500 persons per square mile range. The statistics for the interaction variable indicate that one in six households ("Pop Density >4k and Emp Density >1k", mean=0.16) lie in both the highest three population density ranges and the highest three employment density ranges.

Based on the set of household member variables, the average household contains more than two persons, approximately one worker, more women than men, 1.80 persons age 16 and older, and 0.26 persons age 5 through 15. (Persons younger than 5 were not individually counted in the NHTS.) Of the 1.80 persons age 16 and older, 0.22 of them have a medical condition “making it hard to travel”, 1.01 of them use the internet almost every day, and 0.46 of them never used the internet in the past month. Finally, the average household made two and a half purchases per month on the internet.

Selection of Regression Type



hh-8959.xls

FIGURE 7 Histogram for Dependent Variable (ANNMILES).

Given, as shown in the figure above, that the histogram for the dependent variable (DV) is somewhat similar to a normal curve truncated at zero, several types of regression were considered for the analysis of the above dataset.

First, the Heckman model was considered. The Heckman is for datasets wherein the DV is, at times, not observed. This model may be appropriate for an analysis of *daily* VMT because people who regularly drive do not drive at all on some days, e.g. when they are sick. But this is not the case for analysis of *annual* VMT, as in this dissertation. Households with zero annual VMT do not have unobserved VMT, they simply have zero VMT. Zero-VMT households are similar to low-VMT households: both tend to have few people, have few workers, have low income, and be located in high-proximity areas. For example, for households with more than one person of driving age, one might expect (ceteris paribus) such a household to have multiple vehicles if located in the outer suburbs (and thus high VMT), fewer vehicles (perhaps only one) if located in the inner suburbs (and thus medium VMT), and perhaps zero vehicles if located in the inner city (and thus zero VMT). Therefore, the annual VMT-vs.-proximity (et al.) relationship is essentially a continuous relationship, from high VMT all the way to zero VMT. In fact, as shown in the figure above, as many households have 2000-2999 VMT as have 0 VMT. Therefore, because we do not lack VMT information for a household with zero VMT, the Heckman model is not appropriate for this analysis.

The Tobit model was the second regression type considered. The Tobit model is for datasets wherein certain DV values have been censored. Because no VMT values have been censored, the Tobit model is not appropriate for this analysis.

Given the rejection of the above model candidates, the most widely-used regression type—ordinary least squares (OLS) regression—was selected for the analysis of annual household VMT in the nationwide dataset. Note that the ease of interpreting the coefficients of OLS regressions makes the results of this analysis more readily understood and applied by the target audience of this dissertation.

Regression Analysis

Ordinary least squares models are considered "linear" models in that each independent variable has a linear effect on the dependent variable, in this case VMT, as follows:

$$VMT = \beta_0 + \beta_1 X_1 + \beta_2 X_2 \ldots + \beta_n X_n$$

where $X_{1\text{-}n}$ are the independent variables, $\beta_{1\text{-}n}$ are the coefficients of those independent variables, and β_0 is the "Constant" at the bottom of the regression results.

TABLE 7 VMT-Density OLS Regression Results

Source	SS	df	MS
Model	9.2E+11	32	2.9E+10
Residual	2.1E+12	8926	2.4E+08
Total	3.0E+12	8958	3.4E+08

Number of obs	8,959
F(32, 8926)	121.95
Prob > F	0.0000
R-squared	0.3042
Adj R-squared	0.3017
Root MSE	15,360

DV: ANNMILES	**Coef.**	**Std. Err.**	**t**	**P> |t|**	**Signif***	**95% Conf. Interval**	
Independent Variables- Control							
Basis: HHFAMINC <$20k							
HHFAMINC missing	3,727	745	5.00	0.000	√√	2,267	5,188
HHFAMINC $20,000-$39,999	3,090	564	5.48	0.000	√√	1,984	4,195
HHFAMINC $40,000-$59,999	6,215	616	10.08	0.000	√√	5,007	7,423
HHFAMINC $60,000-$99,999	8,647	630	13.72	0.000	√√	7,412	9,882
HHFAMINC $100,000+	11,245	683	16.46	0.000	√√	9,905	12,584
Home Owned	997	532	1.87	0.061	√	-46	2,041
All Household Members (Age 5+)							
Male Workers (Age 16+)	9,108	426	21.38	0.000	√√	8,273	9,943
Female Workers (Age 16+)	7,200	457	15.76	0.000	√√	6,305	8,096
Male Non-Workers (Age 16+)	4,256	447	9.51	0.000	√√	3,379	5,133
Female Non-Workers (Age 16+)	3,371	460	7.33	0.000	√√	2,469	4,272
Persons Age 5 thru 15	1,057	258	4.10	0.000	√√	551	1,562
Persons 16+ Having MEDCOND	-1,834	390	-4.70	0.000	√√	-2,598	-1,069
Internet Purchases in Past Month	39	32	1.21	0.227	--	-24	102
Persons 16+ Used Internet Almost Every Day	71	320	0.22	0.824	--	-556	698
Persons 16+ Never Used Internet in Past Mo.	-2,140	368	-5.81	0.000	√√	-2,862	-1,419
Basis: Not in MSA or CMSA							
In MSA <1m Persons	527	502	1.05	0.294	--	-458	1,512
In MSA/CMSA >1m Persons	1,173	521	2.25	0.024	√√	152	2,194
Independent Variables- Policy							
Basis: <100 Persons/sqmi, tract							
100-499 Persons/sqmi, tract	-1,705	681	-2.50	0.012	√√	-3,041	-370
500-999 Persons/sqmi, tract	-3,014	985	-3.06	0.002	√√	-4,946	-1,083
1,000-1,999 Persons/sqmi, tract	-2,448	1,022	-2.39	0.017	√√	-4,452	-444
2,000-3,999 Persons/sqmi, tract	-2,771	1,057	-2.62	0.009	√√	-4,843	-698
4,000-9,999 Persons/sqmi, tract	-4,549	1,189	-3.83	0.000	√√	-6,879	-2,218
10,000-24,999 Persons/sqmi, tract	-4,124	1,496	-2.76	0.006	√√	-7,057	-1,191
25,000+ Persons/sqmi, tract	-11,320	1,931	-5.86	0.000	√√	-15,106	-7,533
Basis: <50 Employed Persons/sqmi, tract							
50-99 Employed Persons /sqmi, tract	-399	797	-0.50	0.616	--	-1,961	1,163
100-249 Employed Persons /sqmi, tract	-1,938	798	-2.43	0.015	√√	-3,503	-374
250-499 Employed Persons /sqmi, tract	-2,483	899	-2.76	0.006	√√	-4,245	-720
500-999 Employed Persons /sqmi, tract	-4,189	939	-4.46	0.000	√√	-6,030	-2,348
1,000-1,999 Employed Persons /sqmi, tract	-4,320	1,004	-4.30	0.000	√√	-6,288	-2,352
2,000-3,999 Employed Persons /sqmi, tract	-5,579	1,082	-5.15	0.000	√√	-7,700	-3,457
4,000+ Employed Persons /sqmi, tract	-6,450	1,206	-5.35	0.000	√√	-8,815	-4,086
Pop Density >4k and Emp Density >1k	390	964	0.40	0.686	--	-1,500	2,281
Constant	7,047	802	8.79	0.000	√√	5,475	8,619

* "√": Significant at the 0.10 level; "√√": Significant at the 0.05 level

Tables.xlsx

Prior to discussing the above regression results, the threats to its validity will be addressed.

Threats to Validity

Threats to the validity of the model resulting from the above regression process were checked by addressing the following topics:

- Logical coefficient signs and values
- Influence points
- Normality
- Homoscedasticity
- Linearity
- Independence of error terms
- Model fit
- Self-selection

Logical Coefficient Signs and Values Having examined the signs (i.e. positive vs. negative) of the significant independent variable coefficients, they appear to be logical. For example, the coefficients for each of the five basic person variables [Male Workers (Age 16+), Female Workers (Age 16+), Male Non-Workers (Age 16+), Female Non-Workers (Age 16+), and Persons Age 5 thru 15] are positive, and the coefficient for Persons 16+ Having MEDCOND is negative. Likewise, the values of the coefficient are reasonable. For example, the coefficients for the set of binary income range variables increase with increasing income.

Influence Points Influence points are individual outliers in the data which have an inordinate (and therefore undesirable) impact on the model results. Of the nine scalar independent variables in the model, only one (“Internet Purchases in Past Month”) has a significant range (0-200). But this variable is not significant in the model (t=1.21), eliminating the concern over undue influence from any high values of this variable.

Normality The validity of regression analyses is subject to the normality of the variables involved. According to Hair et al. in their textbook *Multivariate Data Analysis* (*11*):

> "…larger sample sizes reduce the detrimental effects of nonnormality."
>
> "For sample sizes of 200 or more…these same effects [on the results] may be negligible."
>
> "Thus, in most instances, as the sample sizes become large, the researcher can be less concerned about nonnormal variables…."

The sample size of the model (8,959) exceeding 200 observations, the issue of normality was considered not to be problematic.

Homoscedasticity The validity of regression analyses is subject to homoscedasticity, i.e. equal variance of the population error over the range of predictor values. For this analysis, the policy variables (density) being dichotomous (and therefore having no range of values), homoscedasticity is not a concern.

Linearity The validity of the interpretation of this regression analysis is subject to the linearity of the relationship between the policy independent variables (IV) and the dependent variable (DV). The policy IVs in this model (the two sets of density variables) being dichotomous, linearity is not a concern. In fact, the theorized non-linearity of the relationship between proximity and VMT was the purpose of creating the sets of dichotomous density variables.

Independence of Error Terms The validity of regression analyses is subject to the independence of error terms. According to Hair, "we can best identify such an occurrence [independence] by plotting the residuals against any possible sequencing variable" (*11*). Given the use of *annual* VMT for the dependent variable, sequencing (i.e. the date each survey was taken) is not a concern.

Model Fit In addition to the fact that most of the variables in the models (including all but one of the variables in the two sets of density variables) are significantly related to annual VMT (Type I error rate < 0.05), the Adjusted R-squared value is 0.30, demonstrating a good model fit.

Self-Selection Self-selection was addressed in the Data Preparation section above.

Overall Assessment of the Model Given the satisfactory survey of the threats to model validity, it appears that the model is reliable for use in estimating the VMT impact of each level of density.

Regression Results and Findings

The implications of the regression results concerning the control variables will be discussed, followed by the findings concerning the policy variables.

Control Variable Results VMT increases with each rise in income level, as expected. Income appears to have a large impact on VMT, with the highest income being associated with 11,000 miles a year more than that of the lowest income. With a coefficient of approximately 1,000 miles, home ownership has a modest, but statistically significant, *additional* impact on VMT. Given that income is also in the model, the additional VMT impact of home ownership may reflect the higher level of responsibility (resulting in higher credit ratings) which is necessary for both home and auto ownership, and/or it may indicate that the homeowner had higher income in the past (e.g. is retired) which allowed him/her to both buy a home and an auto. Although the driving habits of households in small MSAs do not significantly differ, ceteris paribus, from the base households located outside of MSAs, living in large metro areas—as expected—is associated with modestly higher VMT.

The set of household member variables had expected regression results. Men, ceteris paribus, add more to household VMT than do women, and workers—even controlling for the income effect of working—add twice as much VMT to a household than do non-workers. Children, being too young to drive add modestly to VMT, assumedly due to the additional need for trips created by their presence in the household.

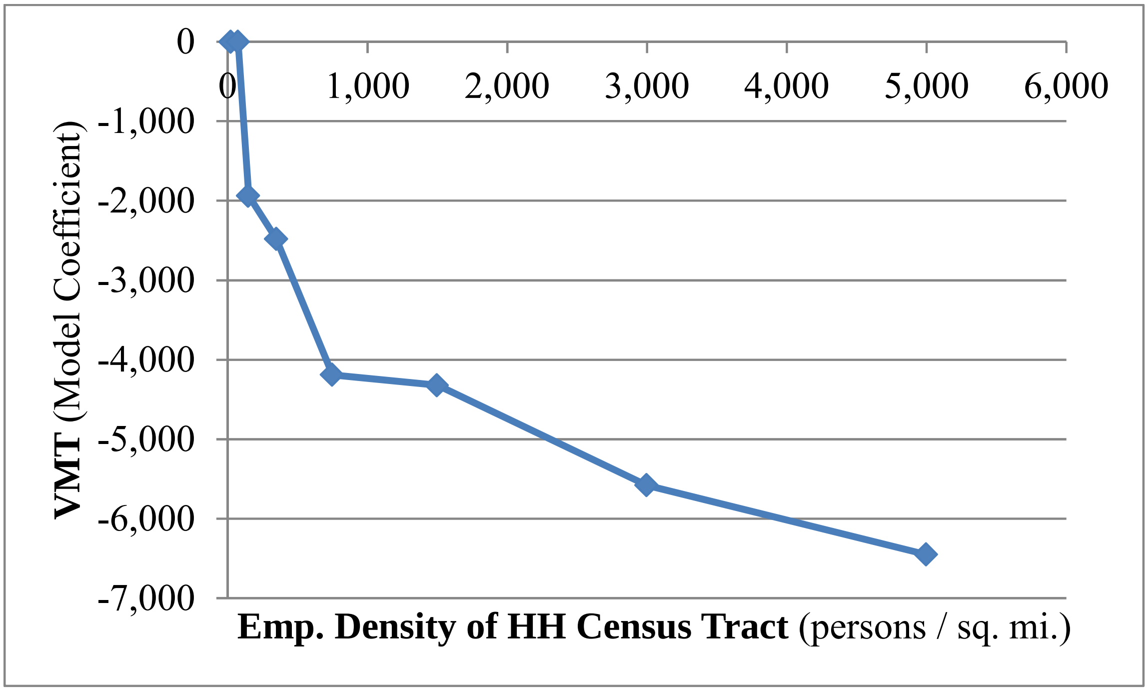
Although the disability variable was highly significant and had the expected negative impact, two of the internet variables—“Internet Purchases in Past Month” and “Persons 16+ Used Internet Almost Every Day”—were not significantly related to VMT at the 0.10 level. The presence of persons who never use the internet had a significant and negative impact on VMT, the negative sign being surprising. This negative relationship may be due to the high age of many of such persons (older people both use the internet less and travel less) and/or the personality type that places persons who are not old in the minority of non-internet usage.

Policy Variable Results and Findings The discussion of density-related findings begins with the lone interaction variable in the regression.

Population and Employment Interaction Results The lack of statistical significance for the “Pop Density >4k and Emp Density >1k” variable (p = 0.686) indicates that this regression provides no evidence of interaction between high population and high employment densities. If this result reflects a real lack of such interaction, then the VMT reduction benefit of a “mixed-use” census tract comes from its high population density and its high employment density, not a synergy between the two. Given 1) the above-reported finding from the literature that census-tract-based proximity has a greater impact on VMT than does neighborhood-based proximity, and 2) the lack of significance of population-employment interaction in this effort’s results, it does not appear that the *interaction* of population and employment in mixed-use developments lowers VMT. In other words, placing much housing and much employment near a household appears to lower the subject household’s VMT, but it may not be necessary to place that housing and employment in the same developments.

VMT vs. Employment Density- Useful Results and Hypothesis Testing Six of the seven variables in the employment variable set being highly significantly related to VMT, their coefficients *fulfill* the research objective—discovering the VMT impact of each level of proximity—and can therefore be used by government to score candidate SGAs according to the expected VMT benefit of their proximity level. .

Although one might see these VMT model results based on census tract density and conclude that *creating* a certain density in a given census tract will give the homes in that census tract the modeled VMT benefit, that conclusion is called into question by the above-stated literature finding that regionally-based proximity has a greater impact on VMT than does neighborhood-based proximity. Although census tracts are larger than neighborhoods, the measured VMT impact of higher density tracts likely reflects the environment *beyond* the subject tract, i.e. its regional environment (in addition to reflecting the environment *within* the subject tract). Therefore, these model results will only be interpreted for providing guidance for identifying desirable census tracts for promoting housing development, not for identifying desirable density levels for census tracts.

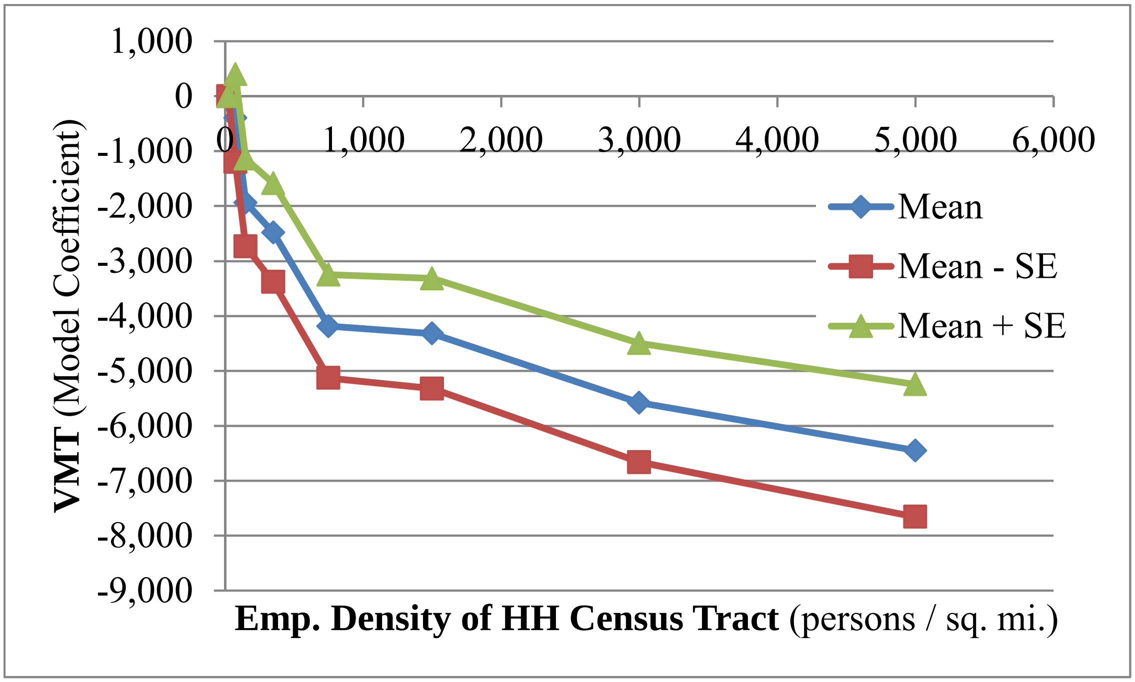


VMT-density curves.xlsx

FIGURE 8 VMT vs. Employment Density of Household Census Tract.

Although the above curve does not exhibit the expected flattening at lower densities, the VMT curve does flatten at the right side as expected and discussed in "The Expected Shape of VMT-Proximity Curves and Key Hypothesis" section above. This flattening provides hope that a sweet spot may be located on the curve.

In preparation of testing the key hypothesis using the employment density variable coefficients shown on the curve, 1) the curve is re-plotted below showing standard errors (SE), and 2) the prevalence of the various NHTS census tract density levels is explored in the table below.



VMT-density curves.xlsx

FIGURE 9 VMT vs. Employment Density of Household Census Tract.

TABLE 8 Prevalence of Employment Density Levels in the U.S.

Employment Density Range, persons/sqmi, tract	**Household Count**	**Weighted Household Count**	**Weighted Household Count, %**	**Weighted Household Count,** percentile range	
<50	2,112	1,296,669	20%	0%	20%
50-99	614	378,456	6%	20%	26%
100-249	1,103	728,345	11%	26%	37%
250-499	1,127	700,887	11%	37%	48%
500-999	1,351	975,329	15%	48%	63%
1,000-1,999	1,284	982,978	15%	63%	78%
2,000-3,999	832	708,769	11%	78%	89%
4,000+	536	716,887	11%	89%	100%
	8,959	**6,488,320**	**100%**		

hh-8959.xlsx

The key hypothesis of this dissertation is:

> There exists a sweet spot on the VMT-proximity curve that has high VMT benefit *and* a proximity level acceptable to many households.

And the specific key hypothesis for testing is:

> The VMT benefit at 67% of maximum proximity is equal to or greater than 80% of the VMT benefit at maximum proximity.

Given that the NHTS labels the highest employment density range as “5,000”, 67% of the maximum proximity level is 3,350 employment per square mile (census tract).

According to the above table, this 3,350 level is approximately the 85 percentile level of U.S. households.

TABLE 9 Hypothesis Testing Worksheet based on Employment Density Curve

Specific Hypothesis: The VMT benefit at 67% of max. proximity is >= 80% of the VMT benefit at max. proximity.
Null Hypothesis: The VMT benefit at 67% of max. proximity is < 80% of the VMT benefit at max. proximity.

			source
Max. proximity:	5,000	employment per sq. mi., census tract	VMT curve
	67%		
67% of max. proximity:	3,350	employment per sq. mi., census tract	
Mean VMT benefit @ 67% of max. prox.:	**5,579**	miles	Regression Table
Mean VMT benefit @ max. prox.:	6,450	miles	Regression Table
	80%		
80% of mean VMT benefit @ max. prox.:	**4,322**	miles	

Therefore, mean VMT benefit at 67% of max. prox. is much higher than 80% of mean VMT benefit at max. prox.

Testing this result considering the standard errors (SE) of the two benefits being compared:

t-test requirements: "two normally distributed but independent populations, σ is unknown" (*10*)
The two populations are mostly independent of each other and σ is unknown.

Difference in the two benefits:	1,258	miles	
SE of VMT benefit @ 67% of max. prox.:	1,082	miles	Regression Table
SE of VMT benefit @ max. prox.:	1,206	miles	Regression Table
	80%		
SE of 80% of VMT benefit @ max. prox.:	965	miles	
Calculated t:	0.87	(calculated via formula for t for comparing two means)	
	vs.		
Critical t value:	1.28	(for α=0.10 and df >1,000)	

Therefore, the null hypothesis is not rejected.

tables.xlsx

Based on the above hypothesis test for the VMT vs. employment density curve:

It is *likely* that the VMT benefit at 67% of maximum proximity is much higher than 80% of the VMT benefit at maximum proximity, but—because the null hypothesis was not rejected—it is not *certain* that the VMT benefit at 67% of maximum proximity is higher than 80% of the VMT benefit at maximum proximity.

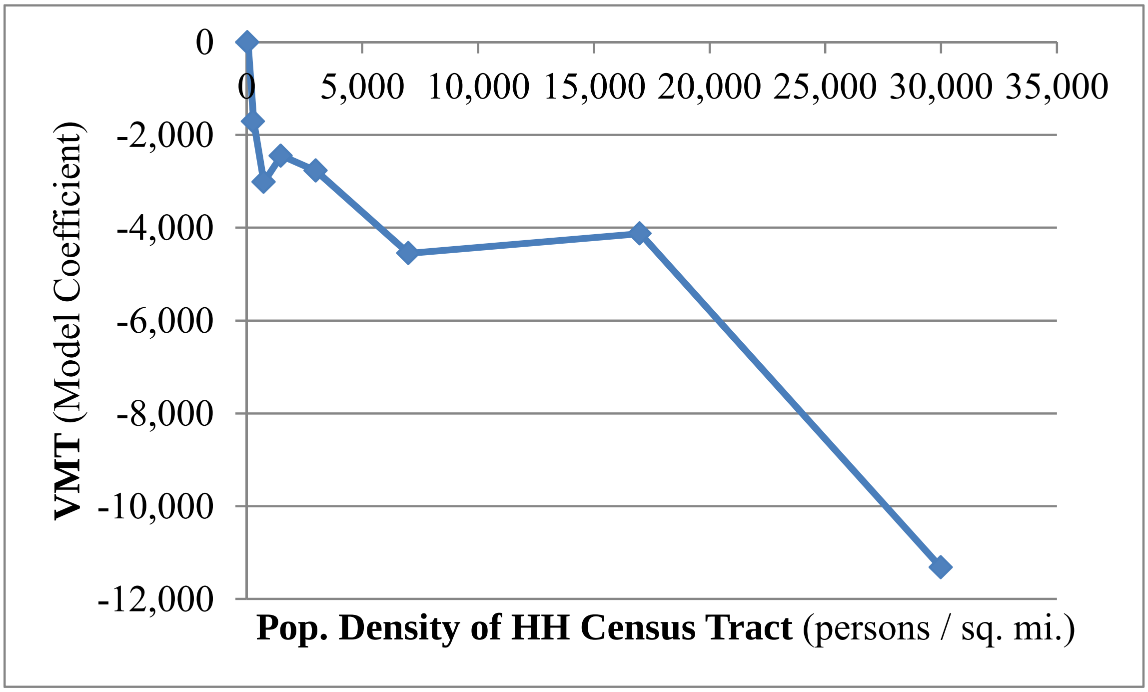
Given a) that the low VMT portion of the above VMT-proximity curve begins at lower proximity than theorized, and b) that the proximity level tested in the hypothesis test falls at the 85 percentile of U.S. households, the curve is examined here for an additional, more moderate point with high VMT benefit and a proximity level acceptable to many households. As an addition to the above hypothesis test of whether the VMT benefit at 67% of maximum proximity is higher than 80% of the VMT benefit at maximum proximity, I also examined whether the VMT benefit at 33% of maximum proximity is higher than 50% of the VMT benefit at maximum proximity.

Given that the NHTS labels the highest employment density range as “5,000”, 33% of the maximum proximity level is 1,650 employment per square mile (census tract). According to the above table, this 1,650 level is approximately the 73 percentile level of U.S. households. Given that a) the mean VMT benefit at this 1,650 level is 4,320 miles, and b) that 50% of the mean VMT benefit at maximum proximity (6,450 * 0.50 = 3,225) is 3,225 miles, i.e. lower than 4,320 miles; the 1,650 level represents an additional, more moderate point on the curve with high VMT benefit and a proximity level acceptable to many households.

Finally, the VMT benefit of this additional point is compared to the VMT benefit of the average household. Given that the employment density VMT benefit of the average U.S. household is 3,115 miles—calculated by weighting the model coefficients according to the above weighted household counts—building new households with the 4,320 mile benefit at this additional point on the curve would lower the average VMT in America.

Given the above analysis of mean VMT benefits—the 1,650 employment per square mile level being acceptable to many households and having a high VMT benefit—it is *likely* that a sweet spot exists on the VMT-proximity curve at the 1,650 employment per square mile level.

VMT vs. Population Density- Useful Results and Hypothesis Testing All seven variables in the population variable set—the lowest of the eight levels being excluded from the regression as the basis for the other seven—being highly significantly related to VMT, their coefficients *fulfill* this dissertation's research objective—discovering the VMT impact of each level of proximity. They can therefore be used by government to score candidate SGAs according to the expected VMT benefit of their proximity level.



VMT-density curves.xlsx

FIGURE 10 VMT vs. Population Density of Household Census Tract.

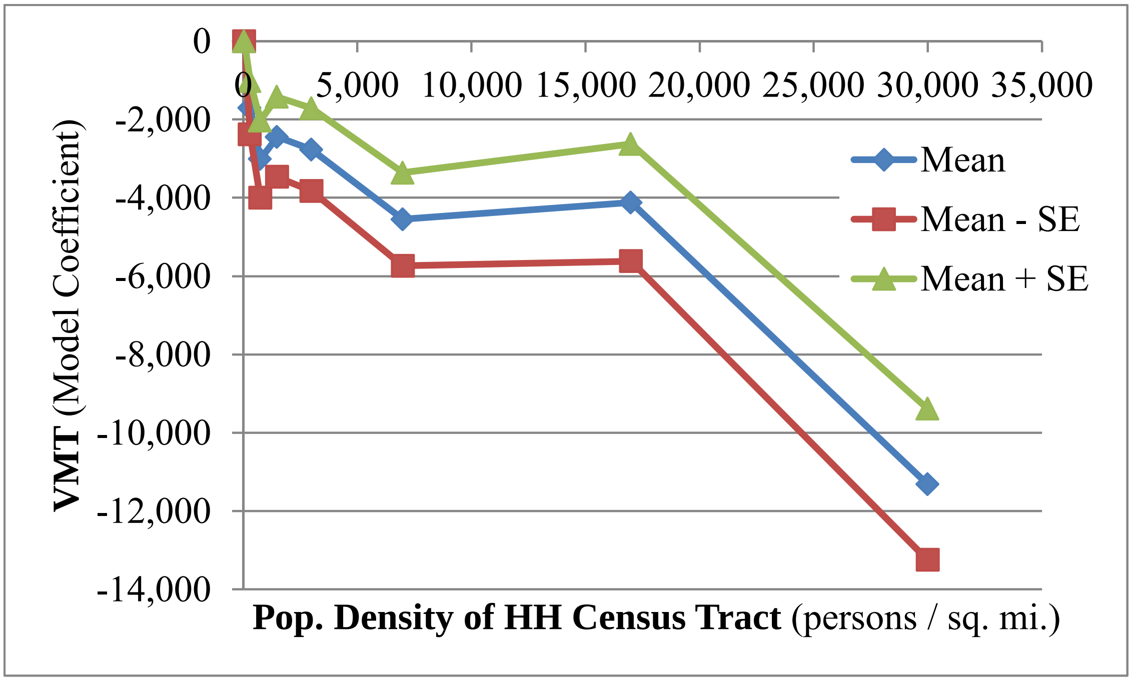
The curve is S-shaped, but not in the fashion expected above in the "The Expected Shape of VMT-Proximity Curve and Secondary Hypothesis" section. Whereas the theoretical curve is flat at low proximities, steep at medium proximities, and flat at high proximities; the empirical curve is steep at low proximities, flat at medium proximities, and steep at high proximities.

The bend in the empirical curve at 10,000-24,999 persons per square mile level (say the 17,500 level) may be related to the aforementioned bend in the public-transit-use-vs.-residential-density curve found by Pushkarev and Zupan (*7*), but given that density was measured in this dissertation at the census tract level, and density was apparently measured at much smaller levels (e.g. block) in the studies reported in Pushkarev and Zupan, it is not possible to conflate the two findings. (Pushkarev and Zupan did not state the level at which density was measured in the synthesis of studies from which they identified the bend in the curve, but—given that they presented the density data in terms of "dwelling units per acre", it appears that these studies measured density over small areas such as blocks.) Tract level population densities are not comparable to block level population densities because—in addition to housing—a census tract will contain land area used for streets and may contain land dedicated to commerce, parks, brownfields, etc. Therefore, a high density block may be found in a low density census tract, and conversely a low density block may be found in a high density census tract.

Likewise, the bend in the VMT-density curve at 17,500 persons per sq. mi. (census tract measure)—above which VMT drops rapidly—may be related to the bend in the auto-ownership-vs.-density curves at 4,500 persons per sq. mi. (zip code measure) found by Dunphy and Fischer (*16*) and Walls et al. (*20*) discussed in the Impetus section above. However, as in the case of Pushkarev and Zupan above, the measurement of density over areas of differing sizes (in this case, census tracts as opposed to zip codes) makes it impossible to conflate the two findings.

After addressing the dramatic drop in VMT associated with the very highest population density level (25,000+ persons per square mile), the curve will be interpreted in detail. Whereas the VMT reduction benefit to a household being situated in moderately high population density census tracts (10,000-24,999 persons per square mile, say the 17,500 level) is approximately 4,000 miles (as compared to being situated in the lowest density tracts, the VMT reduction benefit to a household being situated in *very* high population density census tracts (25,000+ persons per square mile) is approximately 11,000 miles, i.e. almost three times as great. However, given that more than two-thirds of the 126 survey households in this very high density category are located in the New York metro area—which has a level of public transportation system investment much higher than any other metro in the U.S.—it is likely that census tracts with this density either do not exist—or where they do exist, do not have the modeled VMT benefit—in all but the very largest metro areas in the U.S.

Setting aside, therefore, the impact of the eighth and highest population density level, the VMT-population-density curve for the seven "typical" population density levels (i.e. all levels other than the NY-dominated 25,000+ level) exhibits a crescent shape similar to that of the VMT-employment-density curve above. In preparation of testing the key hypothesis using the coefficients for the seven typical population density levels shown on the above curve, 1) the curve is re-plotted below showing standard errors (SE), and 2) the prevalence of the various NHTS census tract density levels is explored in the following table.



VMT-density curves.xlsx

FIGURE 11 VMT vs. Population Density of Household Census Tract.

TABLE 10 Prevalence of Population Density Levels in the U.S.

Population Density Range, persons/sqmi, tract	**Household Count**	**Weighted Household Count**	**Weighted Household Count, %**	**Weighted Household Count,** percentile range	
<100	1,420	967,543	15%	0%	15%
100-499	1,664	960,530	15%	15%	30%
500-999	903	624,780	10%	30%	39%
1,000-1,999	1,225	775,265	12%	39%	51%
2,000-3,999	1,681	1,195,648	18%	51%	70%
4,000-9,999	1,631	1,275,410	20%	70%	89%
10,000-24,999	309	349,808	5%	89%	95%
25,000+	126	339,337	5%	95%	100%
	8,959	**6,488,320**	**100%**		

hh-8959.xlsx

The key hypothesis of this dissertation is:

> There exists a sweet spot on the VMT-proximity curve that has high VMT benefit *and* a proximity level acceptable to many households.

And the specific key hypothesis for testing is:

> The VMT benefit at 67% of maximum proximity is equal to or greater than 80% of the VMT benefit at maximum proximity.

Examining the seven typical population density levels, 67% of the 10,000-24,999 maximum proximity level (say 17,500) is 11,725 persons per square mile (census tract). According to the above table, this 11,725 level is approximately the 90 percentile level of U.S. households.

TABLE 11 Hypothesis Testing Worksheet based on Population Density Curve

Specific Hypothesis: The VMT benefit at 67% of max. proximity is >= 80% of the VMT benefit at max. proximity.
Null Hypothesis: The VMT benefit at 67% of max. proximity is < 80% of the VMT benefit at max. proximity.

			source
Max. proximity (10,000-24,999/sqmi):	17,500	persons per sq. mi., census tract	VMT curve
	67%		
67% of max. proximity:	11,725	persons per sq. mi., census tract	
Mean VMT benefit @ 67% of max. prox.:	**4,549**	miles	Regression Table
(since 11,725 falls in the max. proximity level, use the benefit of the next lowest level, 4,000-9,999 persons/sqmi, tract)			
Mean VMT benefit @ max. prox.:	4,124	miles	Regression Table
	80%		
80% of mean VMT benefit @ max. prox.:	**2,763**	miles	

Therefore, mean VMT benefit at 67% of max. prox. is much higher than 80% of mean VMT benefit at max. prox.

Testing this result considering the standard errors (SE) of the two benefits being compared:

t-test requirements: "two normally distributed but independent populations, σ is unknown" (*10*)
The two populations are mostly independent of each other and σ is unknown.

			source
Difference in the two benefits:	1,786	miles	
SE of VMT benefit @ 67% of max. prox.:	1,189	miles	Regression Table
SE of VMT benefit @ max. prox.:	1,496	miles	Regression Table
	80%		
SE of 80% of VMT benefit @ max. prox.:	1,197	miles	
Calculated t:	1.06	(calculated via formula for t for comparing two means)	
	vs.		
Critical t value:	1.28	(for α=0.10 and df >1,000)	

Therefore, the null hypothesis is not rejected.

tables.xlsx

Based on the above hypothesis test for the VMT vs. population density curve:

It is *likely* that the VMT benefit at 67% of maximum proximity is much higher than 80% of the VMT benefit at maximum proximity, but—because the null hypothesis was not rejected—it is not *certain* that the VMT benefit at 67% of maximum proximity is higher than 80% of the VMT benefit at maximum proximity.

Given a) that the low VMT portion of the above VMT-proximity curve begins at lower proximity than theorized, and b) that the proximity level tested in the hypothesis test falls at the 90 percentile of U.S. households, the curve is examined here for an additional, more moderate point with high VMT benefit and a proximity level acceptable to many households. As an addition to the above hypothesis test of whether the VMT benefit at 67% of maximum proximity is higher than 80% of the VMT benefit at maximum proximity, the question of whether the VMT benefit at 33% of maximum proximity is higher than 50% of the VMT benefit at maximum proximity was also examined.

Given that the highest population density of the seven typical population density levels is 17,500 (10,000-24,999), 33% of the maximum proximity level is 5,775 persons per square mile (census tract). According to the above table, this 5,775 level is approximately the 80 percentile level of U.S. households. Given that a) the mean VMT benefit at this 5,775 level is 4,549 miles, and b) that 50% of the mean VMT benefit at maximum proximity (4,124 * 0.50 = 2,062) is 2,062 miles, i.e. lower than 4,549 miles; the 5,775 level represents an additional, more moderate point on the curve with high VMT benefit and a proximity level acceptable to many households.

Finally, the VMT benefit of this additional point is compared to the VMT benefit of the average household. Given that the population density VMT benefit of the average U.S. household is 3,054 miles—calculated by weighting the model coefficients according to the above weighted household counts—building new households with the 4,549 mile benefit at this additional point on the curve would lower the average VMT in America.

Given the above analysis of mean VMT benefits—the 5,775 persons per square mile level being acceptable to many households and having a high VMT benefit—it is *likely* that a sweet spot exists on the VMT-proximity curve at the 5,775 persons per square mile level.

Effort #2: Identifying Key Proximity Levels for Alternative Modes Using National Data

In order to determine the role played by alternative modes in the VMT-vs.-proximity relationship explored in Effort #1 above, Effort #2 was designed to discover the impact of each proximity level on usage of alternative modes, using national data and density as the measure of proximity, as in Effort #1.

Data Preparation

Choice of Dependent Variable Whereas Effort #1 was conducted using household records, this modal analysis was conducted using person records because different persons in one household may choose different modes. The NHTS variable WRKTRANS was chosen for building the dependent variable because it covers all modes. (PTUSED, for example, only covers public transit.) WKRTRANS records the response to the question: "How did {you/SUBJECT} usually get to work last week?" It should be noted that the use of WRKTRANS limits the analysis to work travel, as opposed to all travel.

Using WRKTRANS, the binary dependent variable "Alternative Mode Used" was created. "Alternative Mode Used" was set equal to 0 if the subject worker used a mode associated by the NHTS with household VMT (i.e. auto, motorcycle), and set equal to 1 if the subject worker used an alternative mode.

Handling Missing Data The 8,959 households in Effort #1 are associated with 18,350 person records in the NHTS person file. Eliminating children (2,306), non-workers (7,900), and work status not attained (3), resulted in 8,141 worker records. Eliminating workers w/o WRKTRANS (1,293 “appropriate skip”, “refused”, and “don’t know”), rendered the 6,848 person records used in Effort #2. As in the case of Effort #1, median densities of population and employment were assumed for the record with missing density data, and missing household income was treated as a category of income.

Data Validity Given that the NHTS (and its predecessor the National Personal Transportation Survey) has been conducted several times (1969, 1977, 1983, 1990, 1995, 2001, and 2009) and is financed by the federal government, the 2009 NHTS data tends to be valid. The “Alternative Mode Used” variable used as the dependent variable in this analysis is based on the respondents’ memory of the primary work mode used “last week”, a naturally reliable response. The key independent variables measuring proximity discussed below (population and employment density by census tract), having been prepared by Nielsen Claritas, are assumed to be reliable.

Although the usage of a robust set of independent variables in this effort’s models removes any requirement that the subject sample dataset reflect exactly the population data, the following table demonstrates the similarity between the weighted full NHTS dataset and the unweighted analysis dataset.

TABLE 12 Similarity Between Full Dataset and Analysis Dataset

		Full Dataset (139,068 workers)				Analysis Dataset (6,848 workers)	
Variable	**NHTS Name**	**Unweighted Count**	**%**	**Weighted Count**	**%**	**Unweighted Count**	**%**
Mode to Work	WRKTRANS						
No Data Available		22,308	16%	21,018,000	14%	0	0%
VMT Mode (auto, motorcyc.)		109,658	79%	117,175,000	77%	6,421	94%
Alternative Mode		7,102	5%	13,179,000	9%	427	6%
		139,068	100%	151,372,000	100%	6,848	100%
Gender	R_SEX						
Male		71,544	51%	81,939,000	54%	3,510	51%
Female		67,524	49%	69,434,000	46%	3,338	49%
		139,068	100%	151,373,000	100%	6,848	100%
Disabled	MEDCOND						
Yes		4,377	3%	4,965,000	3%	179	3%
No, No Data Available		134,691	97%	146,408,000	97%	6,669	97%
		139,068	100%	151,373,000	100%	6,848	100%
Income	HHFAMINC						
HHFAMINC missing		6,751	5%	6,809,000	4%	289	4%
HHFAMINC <$20,000		8,495	6%	15,737,000	10%	365	5%
HHFAMINC $20,000-$39,999		19,885	14%	26,834,000	18%	956	14%
HHFAMINC $40,000-$59,999		23,927	17%	26,215,000	17%	1,180	17%
HHFAMINC $60,000-$99,999		40,249	29%	40,733,000	27%	2,106	31%
HHFAMINC $100,000+		39,761	29%	35,046,000	23%	1,952	29%
		139,068	100%	151,374,000	100%	6,848	100%
Residence	MSASIZE						
No Data Available		1	0%	0	0%	1	0%
Not in MSA or CMSA		26,880	19%	27,974,000	18%	1,359	20%
In MSA <1m Persons		41,582	30%	34,309,000	23%	2,112	31%
In MSA/CMSA >1m Persons		70,605	51%	89,091,000	59%	3,376	49%
		139,068	100%	151,374,000	100%	6,848	100%

Tables.xlsx

Selection and Preparation of Independent Variables Independent variables (IVs) were chosen for this effort's regression based on the theory and literature discussion in the "Preparation" section above. Because VMT is determined by those things which cause one to choose autos over alternative modes, and (if auto has been chosen) those things which affect the annual distance driven, the determinants of usage of alternative modes are largely the same as the determinants of VMT, listed in the table below reproduced from the Preparation section. Therefore, the selection of an IV in the alternative mode regression for each VMT determinant is discussed below.

TABLE 13 Summary of Theorized Determinants of Annual Household VMT

Determinant	Universe
Proximity	Household
Internet Connectivity	Person, Household
Time-Based Accessibility	Household
Public Transit Service Level	Household
Travel Mode Biases ("self-selection")	Person
<u>Socio-economics</u>	
Work Status	Person
Income	Person, Household
Gender	Person
Age	Person
Number of Persons	Household
Disabilities	Person

Tables.xlsx

Proximity As in Effort #1 above, the NHTS density variables based on census tracts—HTEEMPDN for employment density and HTPPOPDN for population density—were chosen to prepare the two sets of density IVs, one for employment and one for population. Because the NHTS variables contain values indicating ranges (e.g. in HTEEMPDN, "75" represents the density range 50-99 employed persons per square mile), a binary variable (e.g. "50-99 Employed Persons /sqmi, tract") was prepared for each range. Because the dataset includes a set of employment variables based on employment locations—the destination of most trips—it is richer than the typical transportation dataset containing only population densities.

In addition—given that the larger the metro area, the greater the distances to destinations—the NHTS variable MSASIZE was used to prepare the IVs "In MSA <1m Persons" and "In MSA/CMSA > 1m Persons" (with basis variable "Not in MSA or CMSA"). Finally, given the lack of statistical significance of the land use mix variable used in Effort #1, no such interaction variable was used in this effort.

Internet Connectivity Given the statistical significance of the "Persons 16+ Never Used Internet in Past Mo." variable in Effort #1 (prepared from the NHTS variable WEBUSE), the binary variable "Never Used Internet in Past Mo." was created from WEBUSE for this person-based analysis. As discussed in Effort #1, the lack of internet usage apparently indicates important travel-related characteristics of the subject person.

Time-based Accessibility, Public Transit Service Level, and Travel Mode Biases Concerning time-based accessibility, public transit service level, and travel mode biases, no NHTS variables were available to directly measure these determinants. Concerning transit service level, however—as discussed in the VMT Theory section above—density is highly related to transit service. Therefore, the impact of transit service on VMT is part of the impact of this effort's density variables, and is measured therefore—along with the other impacts of density—in the coefficients of the density variables. Concerning travel mode biases, these biases (or "self-selection") were addressed in this effort in the Brownstone (*12*) manner discussed in the Preparation section above, i.e. by including several key socio-economic variables in the model.

Socio-economics Concerning socio-economics, all persons in the analysis dataset being workers, “worker status” (listed in the above Determinants table) is moot. Because personal income is not recorded in the NHTS database, household income was used, represented 1) by using the NHTS variable HHFAMINC to prepare the set of binary income IVs (“HHFAMINC $20,000-$39,999”, “HHFAMINC $40,000-$59,999”, etc.), and 2) by using the NHTS variable HOMEOWN to prepare the binary variable “Home Owned.” Gender was represented by using the NHTS variable R_SEX. The dataset covering only workers, age was not injected into the regression analysis. Disabilities were represented by using the NHTS variable MEDCOND to prepare the IV “Persons 16+ Having MEDCOND.”

Descriptive Statistics

As shown in the table below, 6% of the subject workers used alternative modes to get to work. Of those, as many workers walked (157) as used bus and train combined (150).

TABLE 14 Descriptive Statistics- Modal Detail

Mode	Obs	%
VMT Modes		
Auto	6,390	93.3%
Motorcycle	31	0.5%
	6,421	93.8%
Alternative Modes		
Local Public Bus	60	0.9%
Commuter Bus	22	0.3%
Commuter Train	38	0.6%
Subway/Elevated Train	30	0.4%
Bicycle	42	0.6%
Walk	157	2.3%
Other	78	1.1%
	427	6.2%
	6,848	100%

Tables.xlsx

TABLE 15 Descriptive Statistics

	Obs	Mean	Std. Dev.	Min	Max
Dependent Variable					
Alternative Mode Used	6,848	0.06	0.24	0	1
Independent Variables- Control					
Worker Traits					
Male	6,848	0.51	0.50	0	1
Having MEDCOND	6,848	0.03	0.16	0	1
Never Used Internet in Past Mo.	6,848	0.10	0.31	0	1
Household Traits					
Income					
HHFAMINC missing	6,848	0.05	0.22	0	1
HHFAMINC <$20,000	6,848	0.04	0.20	0	1
HHFAMINC $20,000-$39,999	6,848	0.14	0.35	0	1
HHFAMINC $40,000-$59,999	6,848	0.17	0.38	0	1
HHFAMINC $60,000-$99,999	6,848	0.31	0.46	0	1
HHFAMINC $100,000+	6,848	0.29	0.45	0	1
		1.00			
Home Owned	6,848	0.90	0.30	0	1
Size of Metro Area					
Not in MSA or CMSA	6,848	0.20	0.40	0	1
In MSA <1m Persons	6,848	0.31	0.46	0	1
In MSA/CMSA >1m Persons	6,848	0.49	0.50	0	1
		1.00			
Independent Variables- Policy					
Population Density					
<100 Persons/sqmi, tract	6,848	0.15	0.36	0	1
100-499 Persons/sqmi, tract	6,848	0.19	0.39	0	1
500-999 Persons/sqmi, tract	6,848	0.10	0.31	0	1
1,000-1,999 Persons/sqmi, tract	6,848	0.14	0.34	0	1
2,000-3,999 Persons/sqmi, tract	6,848	0.18	0.39	0	1
4,000-9,999 Persons/sqmi, tract	6,848	0.19	0.39	0	1
10,000-24,999 Persons/sqmi, tract	6,848	0.03	0.17	0	1
25,000+ Persons/sqmi, tract	6,848	0.01	0.12	0	1
		1.00			
Employment Density (by location of employment)					
<50 Employed Persons/sqmi, tract	6,848	0.24	0.43	0	1
50-99 Employed Persons /sqmi, tract	6,848	0.07	0.26	0	1
100-249 Employed Persons /sqmi, tract	6,848	0.12	0.33	0	1
250-499 Employed Persons /sqmi, tract	6,848	0.14	0.34	0	1
500-999 Employed Persons /sqmi, tract	6,848	0.15	0.36	0	1
1,000-1,999 Employed Persons /sqmi, tract	6,848	0.14	0.34	0	1
2,000-3,999 Employed Persons /sqmi, tract	6,848	0.09	0.28	0	1
4,000+ Employed Persons /sqmi, tract	6,848	0.05	0.22	0	1
		1.00			

Tables.xlsx

The statistics in 1) the earlier Similarity table (Table 12) and 2) the two Descriptive Statistics tables (Tables 14 and 15) above provide a detailed view of American workers. Concerning the dependent variable, the use of alternative modes is 10% (13,179,000 / [13,179,000+117,175,000]) in the weighted full dataset. Alternative mode usage is 6% (7,102 / [7,102+109,658]) in the unweighted full dataset, and 6% in the unweighted analysis dataset. Alternative mode users were apparently less likely to respond to the NHTS survey.

Concerning the independent variables, the extensive presence of binary variables in the dataset allow for easy categorization of the dataset's households. 51% of the workers in the analysis dataset are male, similar to the male percentage (54%) in the weighted full dataset. The percentage of workers with "medical condition making it hard to travel" (MEDCOND) is 3% in the analysis dataset, the unweighted full dataset, and the weighted full dataset. Half the workers in both unweighted datasets (analysis and full) are located in MSA/CMSAs with more than 1 million population, the other half in less populous areas.

Regression Analysis

Given the binary nature of the dependent variable, logistic regression was performed, with the following results. The coefficients for the independent variables were estimated using the odds value as the dependent measure, as follows, from Hair et al. (*11*):

$$Odds_i = e \wedge (\beta_0 + \beta_1 X_1 + \beta_2 X_2 \ldots + \beta_n X_n)$$

where $Odds_i$ is the odds of using an alternative mode, $X_{1\text{-}n}$ are the independent variables (IVs), $\beta_{1\text{-}n}$ are the coefficients of those IVs, and β_0 is the "Constant" at the bottom of the regression results. Note that the "Odds Ratio", instead of the coefficient, for each IV is reported in the regression results. Given that each IV is binary, the reported "Odds Ratio" indicates the impact of the IV being 1 (or true) on the odds of using an alternative mode. For example, given that the "Odds Ratio" for "Male" is 1.174, being male is associated with a 17% increase in the odds of using an alternative mode. If—based on the values of the other variables—the odds for a female using an alternative mode were 1:6 (or a 14% chance), then the odds for a male using such a mode would be 1.174:6 (1*1.174=1.174), or a 16% chance.

TABLE 16 Alternative-Mode-vs.-Density Logistic Regression Results

Logistic regression		
	Number of obs	6,848
	Log likelihood	-1403
	LR chi^2 (25)	391
	Prob > chi^2	0.0000
	Pseudo R^2	0.122

DV: Alternative Mode Used	**Odds Ratio**	**Std. Err.**	**z**	**P > z**	**Signif***	**95% Conf. Interval**	
Independent Variables- Control							
Worker Traits							
Male	1.174	0.125	1.50	0.13	--	0.952	1.447
Having MEDCOND	1.285	0.359	0.90	0.37	--	0.744	2.222
Never Used Internet in Past Mo.	1.451	0.231	2.34	0.02	√√	1.063	1.982
Household Traits							
Income							
(basis: HHFAMINC <$20k)							
HHFAMINC missing	0.430	0.139	-2.61	0.01	√√	0.228	0.810
HHFAMINC $20,000-$39,999	0.634	0.128	-2.25	0.02	√√	0.426	0.943
HHFAMINC $40,000-$59,999	0.381	0.086	-4.30	0.00	√√	0.245	0.592
HHFAMINC $60,000-$99,999	0.424	0.088	-4.14	0.00	√√	0.282	0.636
HHFAMINC $100,000+	0.545	0.114	-2.90	0.00	√√	0.361	0.821
Home Owned	0.587	0.086	-3.66	0.00	√√	0.441	0.781
Size of Metro Area							
(basis: Not in MSA or CMSA)							
In MSA <1m Persons	0.866	0.161	-0.78	0.44	--	0.602	1.245
In MSA/CMSA >1m Persons	1.160	0.215	0.80	0.42	--	0.807	1.668
Independent Variables- Policy							
Population Density							
(basis: <100 Persons/sqmi, tract)							
100-499 Persons/sqmi, tract	0.858	0.219	-0.60	0.55	--	0.520	1.416
500-999 Persons/sqmi, tract	0.924	0.337	-0.22	0.83	--	0.452	1.890
1,000-1,999 Persons/sqmi, tract	0.456	0.182	-1.97	0.05	√√	0.208	0.997
2,000-3,999 Persons/sqmi, tract	0.695	0.272	-0.93	0.35	--	0.323	1.497
4,000-9,999 Persons/sqmi, tract	0.956	0.380	-0.11	0.91	--	0.439	2.082
10,000-24,999 Persons/sqmi, tract	1.641	0.714	1.14	0.26	--	0.699	3.850
25,000+ Persons/sqmi, tract	7.822	3.626	4.44	0.00	√√	3.152	19.407
Employment Density (by location of employment)							
(basis: <50 Employed Persons/sqmi, tract)							
50-99 Employed Persons /sqmi, tract	0.981	0.299	-0.06	0.95	--	0.539	1.784
100-249 Employed Persons /sqmi, tract	1.266	0.378	0.79	0.43	--	0.705	2.272
250-499 Employed Persons /sqmi, tract	0.867	0.304	-0.41	0.68	--	0.437	1.722
500-999 Employed Persons /sqmi, tract	1.771	0.599	1.69	0.09	√	0.913	3.435
1,000-1,999 Employed Persons /sqmi, tract	1.585	0.560	1.30	0.19	--	0.793	3.167
2,000-3,999 Employed Persons /sqmi, tract	1.792	0.650	1.61	0.11	--	0.880	3.648
4,000+ Employed Persons /sqmi, tract	3.410	1.258	3.33	0.00	√√	1.655	7.029
Constant	0.127	0.033	-8.03	0.00	√√	0.076	0.210

* "√": Significant at the 0.10 level; "√√": Significant at the 0.05 level

Tables.xlsx

Prior to discussing the regression results, the threats to its validity will be addressed.

Threats to Validity

Threats to the validity of the model resulting from the above regression process were checked by addressing the following topics:

- Logical coefficient signs and values
- Influence points
- Normality
- Homoscedasticity
- Linearity
- Independence of error terms
- Model fit
- Self-selection

Logical Coefficient Signs and Values Being a logistic regression, instead of coefficients, odds ratios are published. As do the negative coefficients from which they were calculated, odds ratios less than 1 indicate a negative relationship between the subject IV and the DV, whereas odds ratios greater than 1 indicate a positive relationship. Having examined the odds ratios of the significant independent variable coefficients, they appear to be logical. For example, the odds ratios for all levels of income higher than the low base level are less than 1, as is the odds ratio for having MEDCOND. Likewise, the values of the odds ratios are reasonable when compared between variables. For example, the odds ratios for the significant variables in the two sets of binary density variables increase with increasing density.

Influence Points Influence points are individual outliers in the data which have an inordinate (and therefore undesirable) impact on the model results. All variables used in the regression are binary, eliminating the concern over undue influence from any outlying values.

Normality The validity of regression analyses is subject to the normality of the variables involved. All variables used in the regression are binary, eliminating the concern over any lack of normality.

Homoscedasticity The validity of regression analyses is subject to homoscedasticity, i.e. equal variance of the population error over the range of predictor values. For this analysis, the policy variables (density) being dichotomous (and therefore having no range of values), homoscedasticity is not a concern.

Linearity The validity of the interpretation of this regression analysis is subject to the linearity of the relationship between the policy independent variables (IV) and the dependent variable (DV). The policy IVs in this model (the two sets of density variables) being dichotomous, linearity is not a concern. In fact, the theorized non-linearity of the relationship between proximity and use of alternative transportation was the purpose of creating the sets of dichotomous density variables.

Independence of Error Terms The validity of regression analyses is subject to the independence of error terms. According to Hair, "we can best identify such an occurrence [independence] by plotting the residuals against any possible sequencing variable" (*11*). Although it is expected that time of year affects the choice of alternative modes in colder portions of the US, because the dataset covers the entire United States, sequencing (i.e. the date each survey was taken) is not considered a concern.

Model Fit In addition to the fact that many of the variables in the models—the internet variable, all of the income variables, and two of the binary variables in each of the two sets of density variables—are significantly related to annual VMT (Type I error rate < 0.10), the Pseudo R-squared value is 0.12, demonstrating an adequate model fit.

Self-Selection Self-selection was addressed in the Data Preparation section above.

Overall Assessment of the Model Given the satisfactory survey of the threats to model validity, it appears that the model is reliable for use in estimating the impact of each level of density on the usage of alternative modes.

Regression Results and Findings

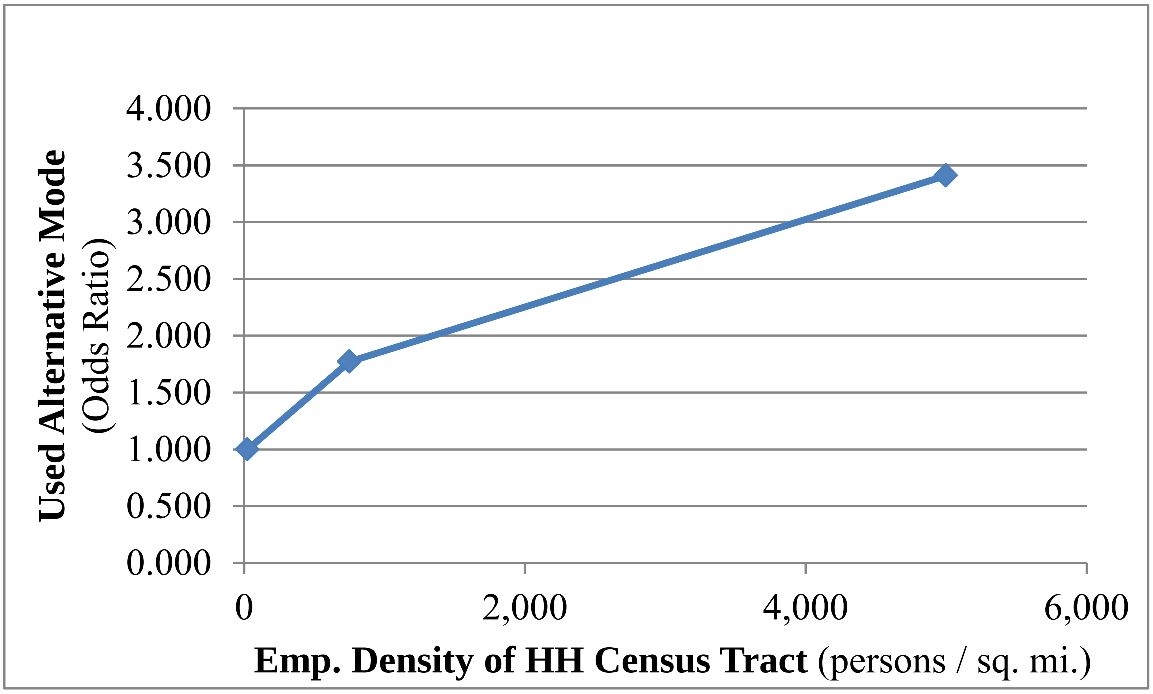
The implications of the regression results concerning the control variables will be discussed, followed by the findings concerning the policy variables.

Control Variable Results Although the odds ratios for all of the variables for higher levels of annual household income ($20k+) being less than 1 is in accordance with stated theory, the odds ratios unexpectedly do not decrease with each rise in income level. Although middle income ($40-$100k) has lower propensity to use alternative transportation than lower income (<$40k), the highest income level ($100k+) surprisingly has a higher odds ratio than that of the middle income levels. Note that this finding of an inconsistent relationship between income and alternative mode usage is mitigated by the fact that the odds ratio of the highest level is within the 95% odds ratio confidence intervals of the middle income levels.

In light of the expectation that women, ceteris paribus, are more likely than men to use public transit, and men are more likely than women to walk; the lack of statistical significance for the gender variable is not surprising. Medical condition and size of metro area are, however, surprisingly statistically insignificant. Although it was expected that home ownership would be related to lower odds of using alternative modes, the strength of that relationship (i.e. the low odds ratio 0.587) is noteworthy given that income is already included in the model.

Policy Variable Results and Findings The discussion of density-related findings begins with employment density and ends with population density.

Alternate Mode vs. Employment Density Results and Findings The odds ratios for the two variables in the employment variable set (of seven total variables) for which the regression resulted in a tight confidence interval (Type I error rate < 0.10) are plotted below.



mode-density curves.xlsx

FIGURE 12 Usage of Alternative Modes for Work vs. Employment Density of Household Census Tract.

Although there were not enough statistically significant binary variables to discover the usage of alternative modes at each employment density level, the results confirm the above-documented literature finding—and the above-stated theory—that usage of alternative modes increases with increasing density, in this case, employment density.

In preparation of the discussion of the implications of these employment density odds ratios, the prevalence of the various NHTS census tract density levels is provided in the table below (as initially shown in Effort #1 above).

TABLE 17 Prevalence of Employment Density Levels in the U.S.

Employment Density Range, persons/sqmi, tract	Household Count	Weighted Household Count	Weighted Household Count, %	Weighted Household Count, percentile range	
<50	2,112	1,296,669	20%	0%	20%
50-99	614	378,456	6%	20%	26%
100-249	1,103	728,345	11%	26%	37%
250-499	1,127	700,887	11%	37%	48%
500-999	1,351	975,329	15%	48%	63%
1,000-1,999	1,284	982,978	15%	63%	78%
2,000-3,999	832	708,769	11%	78%	89%
4,000+	536	716,887	11%	89%	100%
	8,959	**6,488,320**	**100%**		

hh-8959.xlsx

The two statistically significant density odds ratios from the above regression can be discussed in light of the density distribution revealed in the above table. Discussing density levels from lowest to highest, the first level for which the regression produced a statistically significant odds ratio was the 500-999 employment density level. The regression revealed an odds ratio of 1.771 (say 1.8) for this moderate employment density range. This 1.8 odds ratio represents significant alternative mode potential for the 500-999 employment density range. For example, consider a worker who would have 1:24 odds (i.e. 4% chance) of using an alternative mode based on his/her characteristics (e.g. income) and residential location if located in the lowest employment density (<50 persons / sq. mi.). The 1.8 odds ratio indicates that placing that worker's residence in a census tract with 500-999 employment density may increase their odds of using an alternative mode to 1.8:24 (or 1:13), i.e. a 7% chance (1 / [1+13] = 0.07), almost doubling the usage of alternative modes. This doubling may explain a significant portion of the 4,000 annual household VMT benefit associated with this employment density level as revealed in Effort #1 above.

The regression also revealed an odds ratio of 3.410 for the highest (4,000+) employment density range. This odds ratio represents significant alternative mode potential for the 4,000+ employment density range. Comparing this 3.410 to the above 1.771 odds ratio indicates that the odds of using an alternative mode may double (3.410 / 1.771 = 1.93) for a worker with residence in a census tract with the highest employment density, as compared to being located in a census tract of moderate employment density (500-999). This may explain a significant portion of the approximate 2,000 (6,450 - 4,189 = 2,261) annual household VMT benefit revealed by the Effort #1 regression when comparing these two employment density levels.

The usage-of-alternative-modes-vs.-employment-density findings discussed above can be encapsulated as follows:

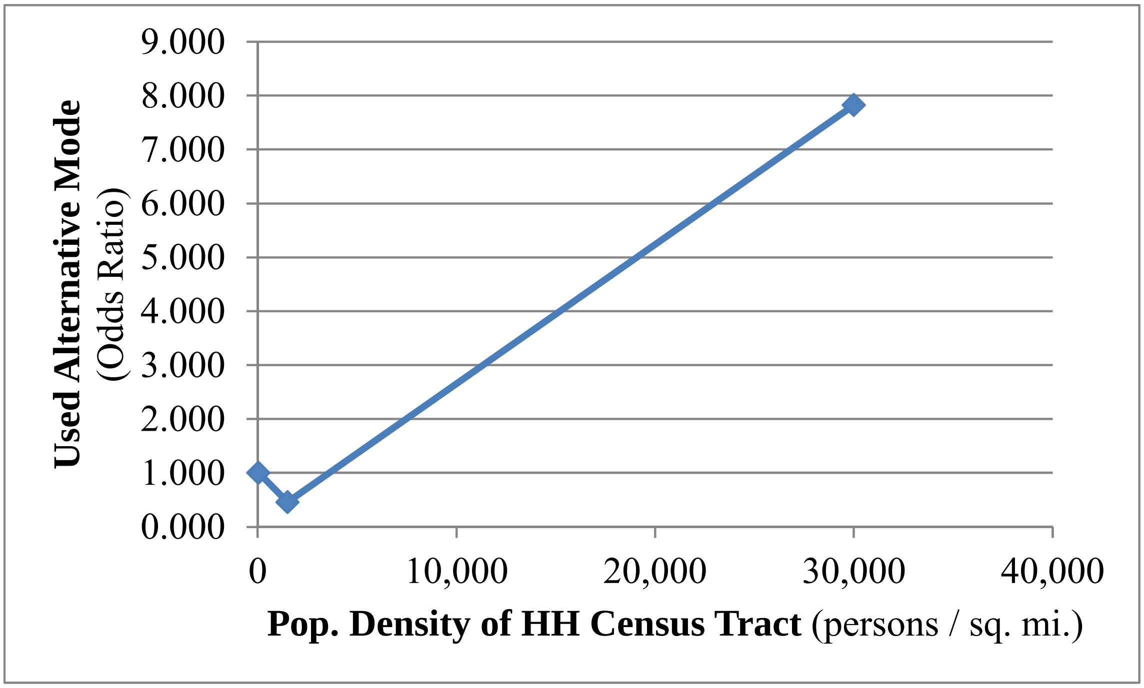
There is an approximate doubling of odds of using an alternative mode (1.8 odds ratio) for a worker with residence situated in census tracts with 500-999 employment per square mile, as compared to being situated in census tracts with the lowest employment density (<50).

This doubling may explain a significant portion of the 4,000 annual household VMT benefit associated with this employment density level.

There is an additional doubling of odds of using an alternative mode (1.93 odds ratio) for a worker with residence situated in census tracts with the highest employment densities (4,000+ employed persons per square mile)—as compared to being situated in census tracts with moderate employment density (500-999 per square mile).

This additional doubling may explain a significant portion of the approximate 2,000 (6,450 - 4,189 = 2,261) annual household VMT benefit when comparing these two employment density levels.

Alternate Mode vs. Population Density Results and Findings The odds ratios for the two variables in the population variable set (of seven total variables) for which the regression resulted in a tight confidence interval (Type I error rate < 0.10) are plotted below, one for the 1,000-1,999 level, and one for the 25,000+ density level.



mode-density curves.xlsx

FIGURE 13 Usage of Alternative Modes for Work vs. Population Density of Household Census Tract.

Although there were not enough statistically significant binary variables to discover the usage of alternative modes at each population density level, the results confirm the above-documented literature finding—and the above-stated theory—that usage of alternative modes increases with increasing density, in this case, population density. The unexpectedly higher usage of alternative modes at the lowest population density level (<100 persons per square mile)—as compared to the moderate density level of 1,000-1,999 persons per square mile—may be due to farmers living at the lowest level reporting “walk” as their mode to work.

No increase in the usage of alternative modes was found at the 500-999 persons per square mile bend in the curve shown on the VMT-vs.-Population-Density curve in Effort #1, indicating that the VMT benefit at that level is a product of shorter driving distances as opposed to greater usage of alternative modes.

Although it would be desirable to compare the results for the statistically significant 1,000-1,999 and 25,000+ person per sq. mi. levels to the aforementioned identification of a bend in the transit-usage-vs.-residential-density curve by Pushkarev and Zupan (*7*), density in the latter work was apparently measured for small areas (e.g. blocks), as opposed to the census tract area densities used in this dissertation. Tract level population densities are not comparable to block level population densities because—in addition to housing—a census tract will contain land area used for streets and may contain land dedicated to commerce, parks, brownfields, etc.; all of which affect census tract density.

In preparation of the discussion of the implications of these population density odds ratios, the prevalence of the various NHTS census tract density levels is provided in the table below (as initially shown in Effort #1 above).

TABLE 18 Prevalence of Population Density Levels in the U.S.

Population Density Range, persons/sqmi, tract	Household Count	Weighted Household Count	Weighted Household Count, %	Weighted Household Count, percentile range	
<100	1,420	967,543	15%	0%	15%
100-499	1,664	960,530	15%	15%	30%
500-999	903	624,780	10%	30%	39%
1,000-1,999	1,225	775,265	12%	39%	51%
2,000-3,999	1,681	1,195,648	18%	51%	70%
4,000-9,999	1,631	1,275,410	20%	70%	89%
10,000-24,999	309	349,808	5%	89%	95%
25,000+	126	339,337	5%	95%	100%
	8,959	**6,488,320**	**100%**		

hh-8959.xlsx

The density odds ratios from the above regression can be discussed in light of the density distribution revealed in the above table. The regression revealed an odds ratio of 7.822 (say 8) for the highest population density range (25,000+ persons per square mile, tract).

This 8 odds ratio represents significant alternative mode potential for the 25,000+ population density range. For example, consider a worker who would have 1:24 odds (i.e. 4% chance) of using an alternative mode based on his/her characteristics (e.g. income) and residential location if located in the lowest population density (<100 persons / sq. mi.). The 8 odds ratio indicates that placing that worker's residence in a census tract with 25,000+ population density may increase their odds of using an alternative mode to 8:24 (or 1:3), i.e. a 25% chance (1 / [1+3] = 0.25), a six-fold increase in the usage of alternative modes. This odds ratio of 8 may explain a significant portion of the 11,000 annual household VMT benefit associated with this population density level as revealed in Effort #1 above.

The alternative-mode-vs.-population-density findings discussed above can be encapsulated as follows:

No increase in the usage of alternative modes was found at the VMT threshold of 500-999 persons per square mile discovered in Effort #1, indicating that the VMT benefit at that level is a product of shorter driving distances as opposed to greater usage of alternative modes.

There is a large increase in the odds of using an alternative mode (8 odds ratio) for a worker with residence situated in census tracts with 25,000+ persons per square mile, as compared to being situated in census tracts with the lowest population density (<100).

This large increase in the propensity to use an alternate mode may explain a significant portion of the 11,000 annual household VMT benefit associated with this highest population density level.

Effort #3: Identifying Key Proximity Levels for VMT Using Hampton Roads Data

Like Effort #1, in order to identify key locations for development, Effort #3 was designed to discover the VMT impact of each level of proximity. Whereas in the national dataset of Effort #1, density was the only available method for measuring proximity, in Effort #3—given the availability of several non-NHTS sources of proximity data for Hampton Roads, Virginia (regional, state, and federal sources)—proximity was measured using a) distance-threshold-based total opportunities, and b) centrality. Adding the new opportunity and centrality data (developed by the author) to the NHTS data created a unique data set.

Data Preparation

The travel and control data for this effort came from the 2009 National Household Travel Survey (NHTS), using the special "DOT" file which contains additional variables not available from the NHTS website. Policy-related opportunity data and centrality data from GIS and the regional four-step model was added to this NHTS data, as described below, to form the final data set.

The data set for this effort was prepared starting with all 3,153 NHTS households located in the thirteen localities represented by the Metropolitan Planning Organization of Hampton Roads, Virginia: Chesapeake, Gloucester, Hampton, Isle of Wight, James City, Newport News, Norfolk, Poquoson, Portsmouth, Suffolk, Virginia Beach, Williamsburg, and York. As in Effort #1 above, the NHTS variable ANNMILES—annual household VMT—was used as the dependent variable.

Handling Missing Data Deleting those households for which the NHTS vehicle file had one or more vehicles with missing ANNMILES (342 households) resulted in 2,811 households (3153-342=2811). Eliminating the 370 households with less than 100% of "household members that completed the interview" (*17*), resulted in 2,441 household records (2811-370 = 2441). Although the usage of a robust set of independent variables in this effort's models removes any requirement that the subject sample dataset reflect exactly the population data, the following table demonstrates the similarity between the weighted full NHTS dataset and the unweighted analysis dataset.

TABLE 19 Similarity Between Full Dataset and Analysis Dataset

		Full Dataset (3,153 HHs)		Analysis Dataset (2,441 HHs)	
Household Variable	**NHTS Name**	**Unweighted Mean**	**Weighted Mean**	**Unweighted Mean**	**Weighted Mean**
Driver Count	DRVRCNT	1.84	1.73	1.77	1.63
Person Count	HHSIZE	2.39	2.49	2.27	2.31
Vehicle Count	HHVEHCNT	2.15	1.93	2.06	1.82
Unit Owned	HOMEOWN	88%	62%	88%	63%
Adult Count	NUMADLT	1.94	1.90	1.84	1.75
Worker Count	WRKCOUNT	0.98	1.11	0.99	1.12

Tables.xlsx

As in Effort #1, missing household income was treated as a category of income, as shown in the "Descriptive Statistics" table below. The three persons age 16+ with missing worker status were assumed not to be workers.

Selection and Preparation of Independent Variables Independent variables (IVs) were chosen for this effort's regression based on the theory and literature discussion in the "Preparation" section above, as summarized in the following table. The selection of an IV for each determinant is discussed below.

TABLE 20 Summary of Theorized Determinants of Annual Household VMT

Determinant	Universe
Proximity	Household
Internet Connectivity	Person, Household
Time-Based Accessibility	Household
Public Transit Service Level	Household
Travel Mode Biases ("self-selection")	Person
<u>Socio-economics</u>	
Work Status	Person
Income	Person, Household
Gender	Person
Age	Person
Number of Persons	Household
Disabilities	Person

Tables.xlsx

Proximity As discussed in the "Measuring Proximity" section above, 1) density, 2) distance-threshold-based total opportunities, and 3) centrality are desirable methods of measuring proximity due to their ease-of-interpretation and theoretical relationship to VMT. Density having been used in Effort #1 above, distance-threshold-based total opportunities and centrality variables were prepared for this third effort. Concerning distance-threshold-based measures, given that both neighborhood and regional proximity have been explored in the literature, in this effort distance-threshold-based opportunity was measured in both the neighborhood and regional environments of the home.

Concerning the neighborhood environment, destinations were measured within one-half mile (Euclidian measure), a walking distance threshold found to be significant in earlier research by the author (*4*). In order to perform measurements at the neighborhood level, the block location of each surveyed household (not being publically available) was obtained from the Virginia Department of Transportation (VDOT). Destinations within one-half mile of each subject household were measured via Geographic Information Systems (GIS) software, creating three variables:

- Retail employees, by place of work, within one-half mile using employment data (2nd Quarter, 2008) from the Virginia Employment Commission (VEC) geo-coded to the street address level.
- Non-retail employees, by place of work, within one-half mile using the same VEC employment data.
- Housing units within one-half mile using Census 2000 data by block.

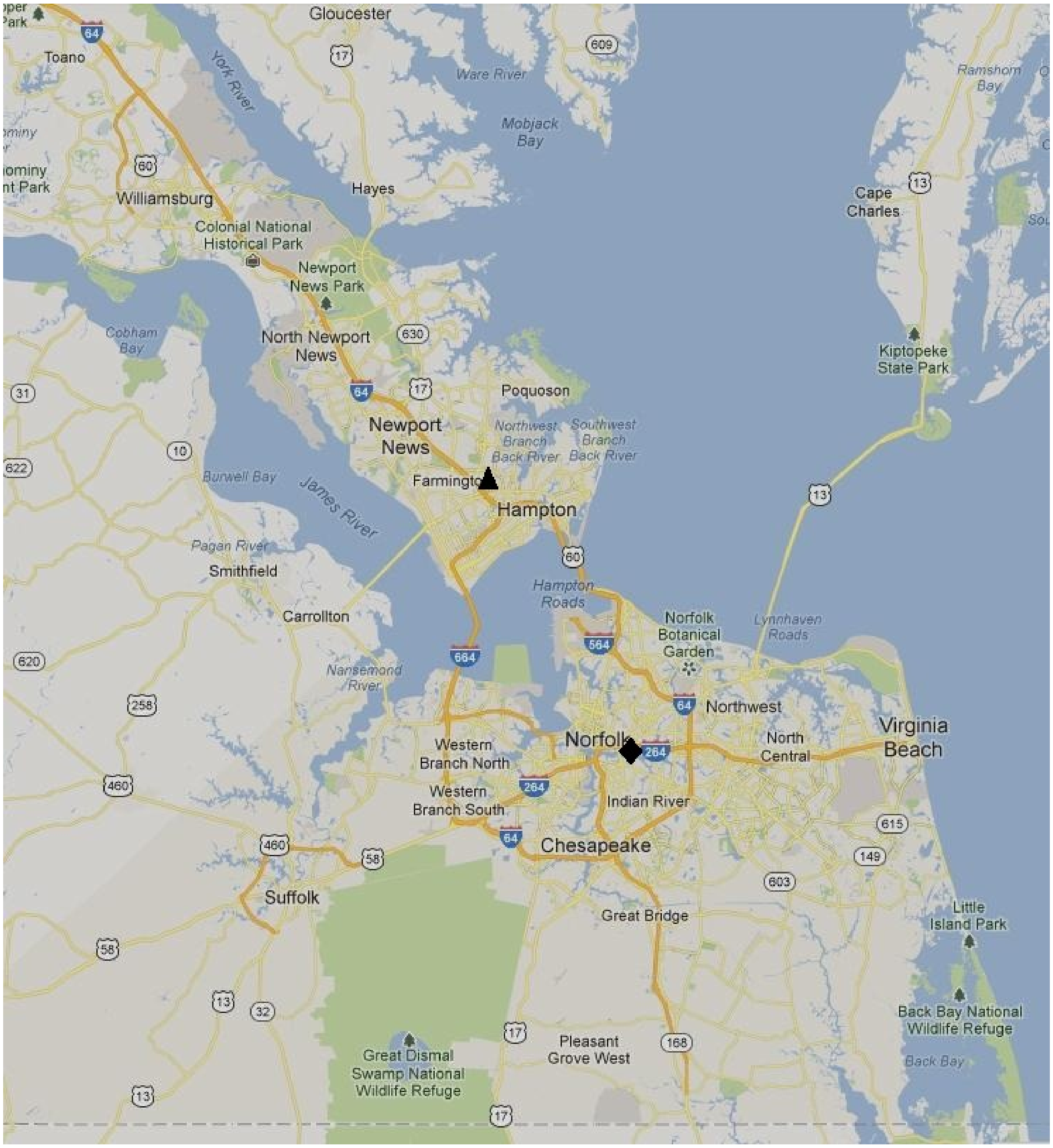
Concerning the regional environment, destinations were typically measured within 10 miles, a threshold found to be significant in exploratory research conducted by the author during the spring semester of 2011 as a foundation for this dissertation. In that research, trip attractions (as calculated by the regional four-step model) were summed within 10, 20, and 40 mile Manhattan distance (i.e. on-street) thresholds for a set of NHTS households in Hampton Roads similar to the set used in this dissertation Effort #3. Regression analysis of the earlier data set revealed that—of the three threshold distances—household VMT is most closely related to destinations within the 10-mile threshold.

In order to reflect the influence of the variety of trip types covered by household VMT, the opportunities in the regional environment of the households in this Effort #3 were measured using three metrics: population, retail employment, and non-retail employment. For population, the 10-mile threshold from the above exploratory research was used. For non-retail employment, based on the exploratory research and Cervero and Duncan's analysis of San Francisco Bay area data which found a median work trip length of 9 miles (*3*), the 10-mile threshold was again used. For retail employment, based on a median shopping trip length of 3 miles in Cervero and Duncan's analysis, a shorter 5-mile threshold was used. Using these distances, regional distance-threshold-based total opportunities variables were prepared for these three metrics—population, non-retail employment, and retail employment—as follows:

- Based on a year 2000 highway network, the on-road distance from the Transportation Analysis Zone (TAZ) of the subject household to each TAZ was measured using a distance table from the regional four-step model obtained from the Hampton Roads Transportation Planning Organization (HRTPO).
- The 2009 population, non-retail employment, and retail employment in each TAZ was obtained from the HRTPO.

In order to be able to plot VMT-opportunity curves, the range of values for each of the three measures were divided into approximately ten sub-ranges, and a binary variable was created for each sub-range. In order that each sub-range represent a statistically valid number of households, the maximum and minimum values of each sub-range were established so that each sub-set would represent roughly 200 households.

Concerning this effort's final proximity measure—centrality—the network distance table discussed above was used to measure the distance from each household to a central point. Because the wide and congested harbor crossings causes people in Hampton Roads to tend to restrict their trips to the side of the harbor in which they live (*28*), a central point was chosen for each side of Hampton Roads—Southside and Peninsula. From an examination of Hampton Roads via Google Maps satellite view, Interstate I-264 & Ballentine Blvd. (represented via a diamond on the map below) appears to be in the middle of Southside activity locations, so it was chosen as the center of the Southside. (This location differing from downtown Norfolk, the conventional "center" of Southside Hampton Roads, the selection was checked using employment data—by employment location—for the subject localities.) Likewise, Peninsula Town Center (represented via triangle on the map below) appears to be in the middle of Peninsula activity locations, so it was chosen as the center of the Peninsula.



hamptonroads.jpg source: Google Maps

FIGURE 14 Hampton Roads, Showing Centers
Legend- triangle: Peninsula Town Center; diamond: I-264 & Ballentine Blvd.

Internet Connectivity Concerning "internet connectivity" in the above table of determinants, the NHTS variable WEBUSE was used to calculate "Persons 16+ Never Used Internet in Past Mo." Because of their lack of significance in the regression in

Effort #1 above, "Persons 16+ Used Internet Almost Every Day" and "Internet Purchases in Past Month" were not included as independent variables in Effort #3.

Time-based Accessibility As discussed in the Preparation section above, the extra access to destinations provided by high-speed roadways can be represented in these models by including a variable measuring the destinations within a certain travel time of the subject household, i.e. "time-based accessibility." As distance-threshold-based total opportunities was chosen to measure proximity, time-threshold-based total opportunities (e.g. population within X minutes) was chosen to measure time-based accessibility.

As in the above case of opportunity, the accessibility of destinations in the regional environment of the households in this Effort #3 were measured using three metrics: population, retail employment, and non-retail employment. In order to have accessibility measures which would work well with the above opportunity measures, the time thresholds of the accessibility measures were calculated based on the distance thresholds of the opportunity measures. For population, assuming 2 minutes per mile for the 10-mile opportunity threshold renders a 20-minute accessibility threshold. Likewise, for non-retail employment, assuming 2 minutes per mile for the 10-mile opportunity threshold renders a 20-minute accessibility threshold. For retail employment, assuming 3 minutes per mile for the 5-mile opportunity threshold (i.e. slower travel due to the expectation of usage of surface streets for these short trips) renders a 15-minute accessibility threshold.

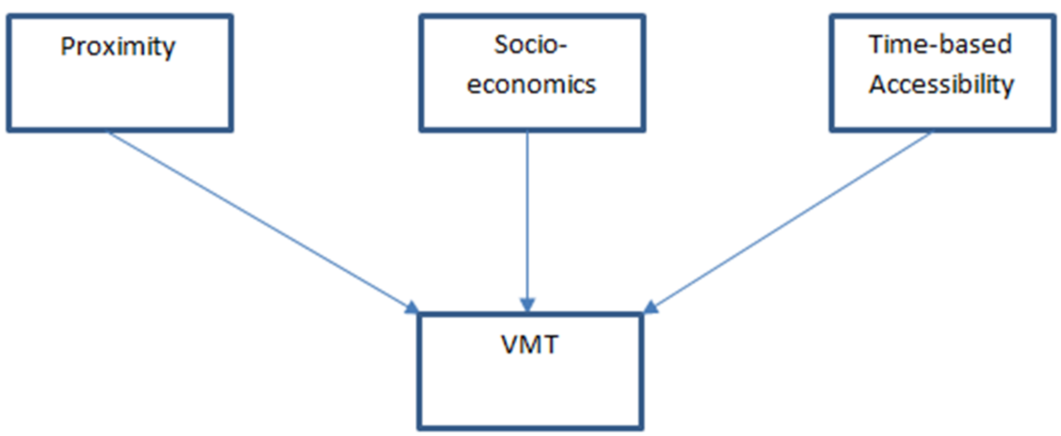
Public Transit Service Level The author conducted exploratory research during the spring semester of 2011 as a foundation for this dissertation using a Hampton Roads

NHTS-based data set similar to that of this effort. He found that only 0.1% of trips were made on public transit. Consequently, public transit service level was not measured for this effort.

Travel Mode Biases ("self-selection") Travel mode biases were addressed in this effort in the Brownstone (12) manner discussed in the Preparation section above, i.e. by including several key socio-economic variables in the model.

Socio-economics As in Effort #1, several key socio-economic variables were included in the models of this effort. From the NHTS data were extracted income, home ownership, work status, gender, age, number of persons, and disabilities.

A drawing of the relationship between the dependent variable and key independent variables is shown below.



key relationships1.png

FIGURE 15 Key Relationships- Effort #3

Data Validity The validity of the above data is a function of the care taken by the agency that collected the original data and the person that processed the data for this analysis (i.e. the author). Given the experience of the agencies that collected the original data—the Federal Highway Administration (FHWA) for the NHTS data, the Census Bureau for the housing units (by block) data, the VEC for the employment (by address) data, and the HRTPO for the distances (by TAZ) and the population and employment (by TAZ) data—the original data is generally trustworthy. Moreover, the fact that the VEC data is collected from unemployment insurance payments lends credibility to the VEC data.

The annual household VMT used as the dependent variable in this analysis is based on the respondents' estimate of annual miles for each household vehicle. Although most people do not know exactly how many miles their vehicles have been driven during the past 12 months, it is expected that the error in those estimates is random and not correlated with any of the independent variables in the analysis.

The key independent variables measuring proximity discussed above (opportunity and centrality), having been prepared by the author, are assumed to be reliable. Although the accuracy of the opportunity variables is limited somewhat by a) the large size of the area units used to measure total opportunities within 10 miles (TAZs), and b) the age of the network used to measure distances (year 2000), the size of TAZs is small compared to 10 miles and few new alignments have been added locally since 2000 which would affect actual year 2009 distances.

TABLE 21 Descriptive Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Dependent Variable					
ANNMILES (annual household VMT)	2,441	18,424	15,664	0	210,000
Independent Variables- Control					
Derived Total Household Income					
basis: HHFAMINC <$20k	2,441	0.10	0.31	0	1
HHFAMINC missing	2,441	0.07	0.26	0	1
HHFAMINC $20,000-$39,999	2,441	0.18	0.38	0	1
HHFAMINC $40,000-$59,999	2,441	0.19	0.39	0	1
HHFAMINC $60,000-$99,999	2,441	0.25	0.43	0	1
HHFAMINC $100,000+	2,441	0.21	0.41	0	1
		1.00			
Home Owned	2,441	0.88	0.32	0	1
All Household Members (Age 5+)					
Male Workers (Age 16+)	2,441	0.52	0.57	0	3
Female Workers (Age 16+)	2,441	0.47	0.57	0	3
Male Non-Workers (Age 16+)	2,441	0.35	0.50	0	2
Female Non-Workers (Age 16+)	2,441	0.55	0.56	0	3
Persons Age 5 thru 15	2,441	0.27	0.67	0	5
Persons 16+ Having MEDCOND	2,441	0.21	0.45	0	2
Persons 16+ Never Used Internet in Past Mo.	2,441	0.41	0.64	0	5
Accessiblity					
2009 Pop. within 20min	2,441	271,852	167,672	5,427	664,761
2009 Retail Emp. within 15min	2,441	14,597	9,592	12	44,193
2009 Other Emp. within 20min	2,441	144,151	101,072	854	414,376
Independent Variables- Policy					
Proximity- Neighborhood Environment					
Retail Emp w/in Half Mile	2,441	86	201	0	3,495
Non-Retail Emp w/in Half Mile	2,441	589	1,080	0	18,809
Housing Units w/in Half Mile	2,441	967	670	0	3,652
Proximity- Regional Environment					
2009 Pop. within 10mi	2,441	270,099	157,690	5,427	639,074
2009 Retail Emp. within 5mi	2,441	10,259	7,186	3	30,698
2009 Other Emp. within 10mi	2,441	150,583	102,038	952	400,550
Distance to Center, mi	2,441	13.78	10.04	0.10	51.85

Tables.xlsx

Descriptive Statistics

The descriptive statistics in the above table provide a detailed view of households in Hampton Roads. Given the difference between weighted and unweighted values in the "Similarity" table above, some of the statistics in the above table of unweighted values will differ from actual average regional values. Concerning the dependent variable, the average household VMT is 18,424, similar to the 19,011 figure from the national dataset used in Effort #1. Concerning the independent variables, the usage of a set of binary variables to represent household income allows for easy categorization of the dataset's households. With approximately half of the households in the lower three income levels and an equal share in the higher two income levels, median household income is approximately $60,000, somewhat higher than the $50,000 of the national dataset. Fortunately, only 7% of the household records are missing income information. Although home ownership in the dataset is very high (88%), note that the weighted value shown in the "Similarity" table above is significantly lower (63%).

The surveyed households represent a broad range of (distance-threshold-based) opportunity, centrality, and accessibility values. Concerning the neighborhood environment, households ranged from the rural condition of having zero employment and housing units within one-half mile to the urban condition of having thousands of employment and housing units within that distance. Concerning centrality, surveyed households were located in a range of less than one mile to more than fifty miles from the subject metro center.

Based on the set of household member variables, the average surveyed household contains more than two persons, approximately one worker, more women than men, 1.89

persons age 16 and older, and 0.27 persons age 5 through 15. (Persons younger than 5 were not individually counted in the NHTS.) Of the 1.89 persons age 16 and older, 0.21 of them have a medical condition “making it hard to travel”, and 0.41 of them never used the internet in the past month.

VMT vs. Opportunity

Regression Analysis As in Effort #1 where the VMT impact of each level of density was determined, to determine the VMT impact of each level of opportunity in Effort #3, OLS regression was used. Initially, an OLS regression was run using all of the theory-based variables prepared as discussed above representing socio-economics, internet connectivity, high-speed roadways (represented via three time-threshold-based accessibility variables), and opportunity (represented via four sets of distance-threshold-based total opportunities variables). (As discussed below, another regression was run later that excluded the insignificant variables of the initial run.) The results of the initial regression are shown below.

TABLE 22 Initial VMT-Opportunity OLS Regression Results (page one)

Source	SS	df	MS		
Model	2.5E+11	45	5.6E+09	**Number of obs**	2,441
Residual	3.5E+11	2395	1.5E+08	**F(32, 8926)**	38.16
Total	6.0E+11	2440	2.5E+08	**Prob > F**	0.0000
				R-squared	0.4176
				Adj R-squared	0.4066
				Root MSE	12,068

DV: ANNMILES	Coef.	Std. Err.	t	P> |t|	Signif*	95% Conf. Interval	
Control Variables							
Household Family Income							
Basis: HHFAMINC <$20k							
HHFAMINC missing	1,769	1,210	1.46	0.144	--	-604	4,142
HHFAMINC $20,000-$39,999	1,493	982	1.52	0.129	--	-434	3,419
HHFAMINC $40,000-$59,999	3,063	1,029	2.98	0.003	√√	1,046	5,081
HHFAMINC $60,000-$99,999	4,877	1,042	4.68	0.000	√√	2,833	6,921
HHFAMINC $100,000+	8,486	1,129	7.51	0.000	√√	6,272	10,701
Home Owned	2,925	818	3.58	0.000	√√	1,321	4,528
All Household Members (Age 5+)							
Male Workers (Age 16+)	10,950	573	19.11	0.000	√√	9,826	12,073
Female Workers (Age 16+)	8,491	576	14.73	0.000	√√	7,361	9,622
Male Non-Workers (Age 16+)	6,010	597	10.07	0.000	√√	4,840	7,181
Female Non-Workers (Age 16+)	5,387	598	9.01	0.000	√√	4,215	6,560
Persons Age 5 thru 15	585	392	1.49	0.135	--	-183	1,353
Persons 16+ Having MEDCOND	-2,145	589	-3.64	0.000	√√	-3,301	-990
Persons 16+ Never Used Internet in Past Mo.	-2,243	468	-4.79	0.000	√√	-3,161	-1,326
Accessibility							
2009 Pop. within 20min	0.015	0.010	1.42	0.155	--	-0.005	0.034
2009 Retail Emp. within 15min	-0.030	0.091	-0.33	0.740	--	-0.209	0.148
2009 Other Emp. within 20min	-0.028	0.015	-1.92	0.054	√	-0.058	0.001

* "√": Significant at the 0.10 level; "√√": Significant at the 0.05 level

TABLE 22 Initial VMT-Opportunity OLS Regression Results (page two)

DV: ANNMILES	Coef.	Std. Err.	t	P> |t|	Signif*	95% Conf. Interval	
Policy Variables							
Neighborhood Environment							
Retail Emp w/in Half Mile	0.405	1.378	0.29	0.769	--	-2.297	3.107
Non-Retail Emp w/in Half Mile	0.184	0.262	0.70	0.482	--	-0.330	0.699
Housing Units w/in Half Mile	0.060	0.533	0.11	0.910	--	-0.984	1.104
Regional Population Environment							
Basis: 2009 Pop. within 10mi, 0-50k							
2009 Pop. within 10mi, 50k-100k	-4,385	2,161	-2.03	0.043	√√	-8,622	-147
2009 Pop. within 10mi, 100k-200k	-6,335	2,626	-2.41	0.016	√√	-11,485	-1,185
2009 Pop. within 10mi, 200k-250k	-6,962	2,980	-2.34	0.020	√√	-12,807	-1,117
2009 Pop. within 10mi, 250k-300k	-7,289	3,177	-2.29	0.022	√√	-13,520	-1,059
2009 Pop. within 10mi, 300k-350k	-7,232	3,551	-2.04	0.042	√√	-14,195	-269
2009 Pop. within 10mi, 350k-450k	-6,278	4,088	-1.54	0.125	--	-14,293	1,738
2009 Pop. within 10mi, 450k-550k	-8,512	4,608	-1.85	0.065	√	-17,548	524
2009 Pop. within 10mi, 550k+	-8,599	5,116	-1.68	0.093	√	-18,631	1,432
Regional Retail Environment							
Basis: 2009 Ret. Emp. w/in 5mi, 0-1.5k							
2009 Ret. Emp. w/in 5mi, 1.5k-3k	-72	1,156	-0.06	0.950	--	-2,340	2,196
2009 Ret. Emp. w/in 5mi, 3k-5k	-913	1,412	-0.65	0.518	--	-3,682	1,856
2009 Ret. Emp. w/in 5mi, 5k-7.5k	-301	1,474	-0.20	0.838	--	-3,191	2,589
2009 Ret. Emp. w/in 5mi, 7.5k-10k	-2,731	1,679	-1.63	0.104	--	-6,024	562
2009 Ret. Emp. w/in 5mi, 10k-12.5k	-1,060	1,718	-0.62	0.537	--	-4,429	2,309
2009 Ret. Emp. w/in 5mi, 12.5k-15k	-1,212	1,932	-0.63	0.530	--	-5,000	2,576
2009 Ret. Emp. w/in 5mi, 15k-17.5k	-1,141	2,132	-0.54	0.593	--	-5,321	3,040
2009 Ret. Emp. w/in 5mi, 17.5k-22.5k	-513	2,347	-0.22	0.827	--	-5,116	4,090
2009 Ret. Emp. w/in 5mi, 22.5k+	-881	2,818	-0.31	0.754	--	-6,408	4,645
Regional Other Employment Environment							
Basis: 2009 Oth. Emp. w/in 10mi, 0-20k							
2009 Oth. Emp. w/in 10mi, 20k-50k	2,143	2,180	0.98	0.326	--	-2,132	6,418
2009 Oth. Emp. w/in 10mi, 50k-80k	2,360	2,571	0.92	0.359	--	-2,682	7,402
2009 Oth. Emp. w/in 10mi, 80k-110k	3,702	2,843	1.30	0.193	--	-1,873	9,276
2009 Oth. Emp. w/in 10mi, 110k-140k	2,229	3,039	0.73	0.464	--	-3,732	8,189
2009 Oth. Emp. w/in 10mi, 140k-150k	1,775	3,181	0.56	0.577	--	-4,462	8,013
2009 Oth. Emp. w/in 10mi, 150k-180k	3,244	3,181	1.02	0.308	--	-2,994	9,482
2009 Oth. Emp. w/in 10mi, 180k-250k	2,383	3,549	0.67	0.502	--	-4,577	9,343
2009 Oth. Emp. w/in 10mi, 250k-310k	3,723	3,863	0.96	0.335	--	-3,852	11,298
2009 Oth. Emp. w/in 10mi, 310k+	3,756	4,263	0.88	0.378	--	-4,604	12,115
Constant	3,216	1,475	2.18	0.029	√√	323	6,109

* "√": Significant at the 0.10 level; "√√": Significant at the 0.05 level

Tables.xlsx

With an R-squared value exceeding 0.4, the above regression shows an excellent statistical relationship between the theory-based set of independent variables and annual household VMT. Concerning the control variables, most are statistically significant and have coefficients with logical size and sign. The socio-economic variables are highly significant, and the Accessibility set of variables has mixed statistical significance—

population accessibility having moderate significance, retail accessibility having practically no significance, and other employment accessibility having high significance. Concerning the policy variables, the Regional *Population* Environment set of opportunity variables is highly statistically significant, but the Regional *Retail* Environment, and Regional *Other Employment* Environment sets of variables have practically no statistical significance. The insignificance of employment-based opportunity in this effort, contrasts with the significance of employment density in Effort #1. Finally, the Neighborhood Environment set of variables has little statistical significance. This latter result concurs with the above-reported research of Bagley and Mokhtarian (*1*) who found "little…effect of neighborhood type on VMT…"

Given the VMT significance of regionally-based opportunity, and the lack of VMT significance of neighborhood-based opportunity, demonstrated by this effort and the literature,

Effort #3 identifies the best *regional locations* for the promotion of housing development, and does not address the best *neighborhood form* of the housing to be built at those locations.

The housing which is built in VMT-desirable locations (desirable due to its regional location) may itself consist of one or many units, and single-family or multi-family units—these choices depending on the availability of land, market demand, and the ultimate design of the subject city favored by government. Likewise, Effort #1 identifies the best census tracts for the promotion of housing development, and does not address the best neighborhood form of the housing to be built in those tracts. Although it is widely held that neighborhood form affects the usage of alternative modes, it was not necessary to address neighborhood form in Effort #2 because its purpose was to parse the findings of Effort #1.

In order to prevent the large number of insignificant variables from affecting the coefficients of the (statistically significant) Regional Population Environment set of variables; the Regional Retail, Regional Other Employment, and Neighborhood Environment sets of variables were removed for the final regression. In order to reflect the impact of high-speed roadways, the population accessibility variable was retained to be paired with the population set of opportunity variables. Because the Regional Retail and Regional Other Employment variable sets were dropped, the Retail Employment and Other Employment accessibility variables were also dropped. (See "Time-based Accessibility" above for discussion of pairing accessibility with opportunity to allow accessibility to reflect the impact of high-speed roadways on VMT.)

The resulting final VMT-opportunity regression (with opportunity measured via the distance-threshold-based total opportunities measure of population within 10 miles) is shown on the following page:

TABLE 23 Final VMT-Opportunity OLS Regression Results

Source	SS	df	MS
Model	2.5E+11	22	1.1E+10
Residual	3.5E+11	2418	1.5E+08
Total	6.0E+11	2440	2.5E+08

Number of obs	2,441
F(32, 8926)	77.29
Prob > F	0.0000
R-squared	0.4129
Adj R-squared	0.4075
Root MSE	12,059

DV: ANNMILES	Coef.	Std. Err.	t	P> |t|	Signif*	95% Conf. Interval	
Control Variables							
Household Family Income							
Basis: HHFAMINC <$20k							
HHFAMINC missing	1,957	1,202	1.63	0.103	--	-399	4,313
HHFAMINC $20,000-$39,999	1,656	977	1.70	0.090	√	-259	3,571
HHFAMINC $40,000-$59,999	3,060	1,022	2.99	0.003	√√	1,056	5,064
HHFAMINC $60,000-$99,999	5,010	1,033	4.85	0.000	√√	2,984	7,036
HHFAMINC $100,000+	8,693	1,113	7.81	0.000	√√	6,510	10,877
Home Owned	3,100	806	3.85	0.000	√√	1,519	4,680
All Household Members (Age 5+)							
Male Workers (Age 16+)	10,951	568	19.27	0.000	√√	9,837	12,065
Female Workers (Age 16+)	8,519	569	14.96	0.000	√√	7,402	9,635
Male Non-Workers (Age 16+)	5,997	593	10.12	0.000	√√	4,835	7,159
Female Non-Workers (Age 16+)	5,485	592	9.26	0.000	√√	4,323	6,646
Persons Age 5 thru 15	631	387	1.63	0.103	--	-128	1,390
Persons 16+ Having MEDCOND	-2,153	587	-3.67	0.000	√√	-3,304	-1,002
Persons 16+ Never Used Internet in Past Mo.	-2,310	465	-4.97	0.000	√√	-3,222	-1,399
Accessibility							
2009 Pop. within 20min	-0.002	0.005	-0.38	0.703	--	-0.011	0.007
Policy Variables							
Regional Population Environment							
Basis: 2009 Pop. within 10mi, 0-50k							
2009 Pop. within 10mi, 50k-100k	-3,545	1,128	-3.14	0.002	√√	-5,757	-1,332
2009 Pop. within 10mi, 100k-200k	-3,911	1,267	-3.09	0.002	√√	-6,396	-1,426
2009 Pop. within 10mi, 200k-250k	-4,861	1,420	-3.42	0.001	√√	-7,645	-2,076
2009 Pop. within 10mi, 250k-300k	-5,553	1,556	-3.57	0.000	√√	-8,604	-2,501
2009 Pop. within 10mi, 300k-350k	-5,612	1,747	-3.21	0.001	√√	-9,036	-2,187
2009 Pop. within 10mi, 350k-450k	-4,772	2,084	-2.29	0.022	√√	-8,858	-685
2009 Pop. within 10mi, 450k-550k	-6,470	2,486	-2.60	0.009	√√	-11,345	-1,596
2009 Pop. within 10mi, 550k+	-6,473	3,009	-2.15	0.032	√√	-12,373	-573
Constant	3,127	1,357	2.30	0.021	√√	467	5,788

* "√": Significant at the 0.10 level; "√√": Significant at the 0.05 level

Tables.xlsx

Prior to discussing the above final VMT-opportunity regression results, the threats to its validity will be addressed.

Threats to Validity Threats to the validity of the model resulting from the above regression process were checked by addressing the following topics:

- Logical coefficient signs and values
- Influence points
- Normality
- Homoscedasticity
- Linearity
- Independence of error terms
- Model fit
- Self-selection

Logical Coefficient Signs and Values Having examined the signs (i.e. positive vs. negative) of the significant independent variable coefficients, they appear to be logical. For example, the coefficients for each of the five basic person variables [Male Workers (Age 16+), Female Workers (Age 16+), Male Non-Workers (Age 16+), Female Non-Workers (Age 16+), and Persons Age 5 thru 15] are positive, and the coefficient for Persons 16+ Having MEDCOND is negative. Likewise, the values of the coefficient are reasonable. For example, the coefficients for the set of binary income range variables increase with increasing income.

Influence Points Influence points are individual outliers in the data which have an inordinate (and therefore undesirable) impact on the model results. Of the eight scalar independent variables in the model, seven count the number of persons of a certain type in the household. The maximum value of all seven variables being 5, there are no outliers. The final scalar variable (population within 20 minutes) has a significant range (5,427-664,761), but an examination of a histogram of this variable reveals no outliers, eliminating the concern over undue influence from stray low or high values of this variable.

Normality The validity of regression analyses is subject to the normality of the variables involved. According to Hair et al. in their textbook *Multivariate Data Analysis* (*11*):

> "…larger sample sizes reduce the detrimental effects of nonnormality."
>
> "For sample sizes of 200 or more…these same effects [on the results] may be negligible."
>
> "Thus, in most instances, as the sample sizes become large, the researcher can be less concerned about nonnormal variables…."

The sample size of the model (2,441) exceeding 200 observations, the issue of normality was considered not to be problematic.

Homoscedasticity The validity of regression analyses is subject to homoscedasticity, i.e. equal variance of the population error over the range of predictor values. For this analysis, the set of policy variables (Regional Population Environment) being dichotomous and therefore having no range of values, homoscedasticity is not a concern.

Linearity The validity of the interpretation of this regression analysis is subject to the linearity of the relationship between the policy independent variables (IV) and the dependent variable (DV). The policy IVs in this model being dichotomous, linearity is not a concern. In fact, the theorized non-linearity of the relationship between proximity and VMT was the purpose of creating the set of dichotomous opportunity variables.

Independence of Error Terms The validity of regression analyses is subject to the independence of error terms. According to Hair, "we can best identify such an occurrence [independence] by plotting the residuals against any possible sequencing variable" (*11*). Given the use of *annual* VMT for the dependent variable, sequencing (i.e. the date each survey was taken) is not a concern.

Model Fit In addition to the fact that most of the variables in the models (including all of the policy variables) are significantly related to annual VMT (Type I error rate < 0.05), the Adjusted R-squared value is 0.41, demonstrating an excellent model fit.

Self-Selection Self-selection was addressed in the Data Preparation section above.

Overall Assessment of the Model Given the satisfactory survey of the threats to model validity, it appears that the model is reliable for use in estimating VMT impact by opportunity level.

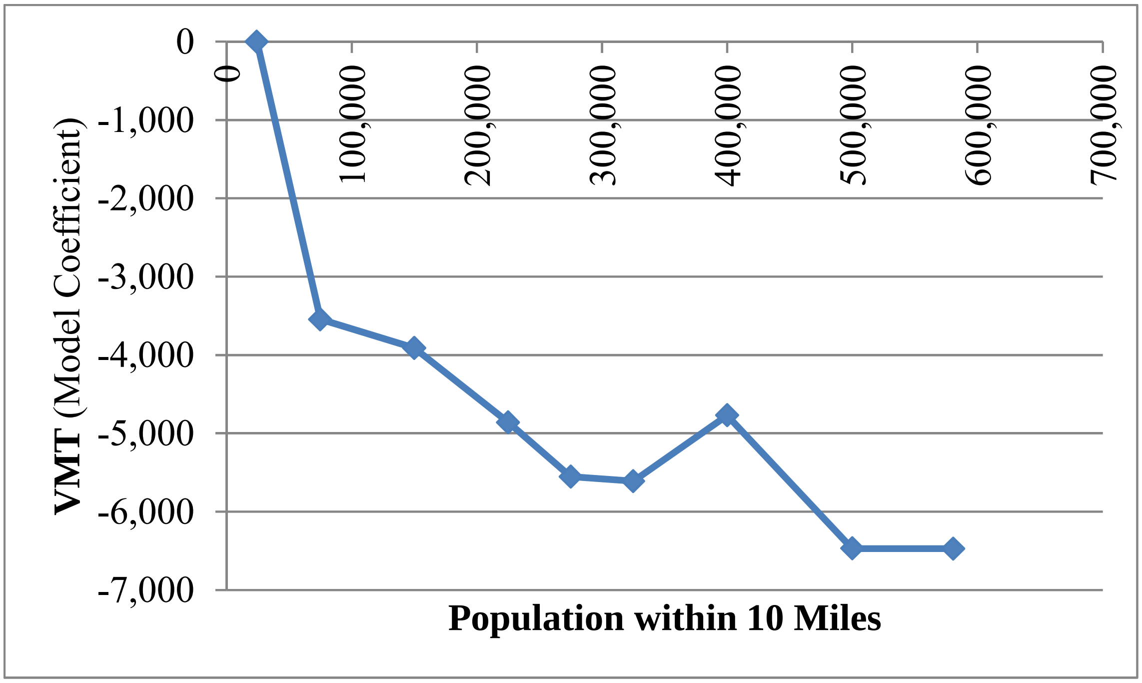
Useful Regression Results and Hypothesis Testing The results of the regression concerning the control variables will be discussed first, followed by the results concerning the policy variables.

Control Variables First, most of the control variables behaved as expected. VMT increases with each rise in income level. And income appears to have a large impact on VMT, with the highest income being associated with approximately 9,000 miles a year more than that of low income, a similar result to that of Effort #1 (11,000 miles). Home ownership has a large *additional* impact on VMT (3,000 miles), three times that of the home ownership variable in Effort #1 (1,000).

The set of household member variables had expected regression results. Men, ceteris paribus, add more to household VMT than do women, and workers—even controlling for the income effect of working—add almost twice as much VMT to a household than do non-workers. Children, being too young to drive, add only modestly to VMT. The disability variable was highly significant and had the expected negative impact, and the presence of persons who never use the internet had a significant and negative impact on VMT. As in Effort #1, this negative relationship may be due to the high age of many of such persons (older people both use the internet less and travel less) and/or the personality type that places persons who are not old in the minority of non-internet usage.

The final control variable, the accessibility variable “2009 Pop. within 20min” intended to reflect the impact of high-speed roadways available to the subject household—was highly insignificant (Type I error rate = 0.703). This may indicate that high-speed roadways do not contribute as much to VMT as theorized above.

Policy Variables- Useful Results and Hypothesis Testing It should be noted that—because 1) population within 10 miles is highly correlated to employment within 10 miles, and 2) the two employment-based opportunity variable sets dropped out of the regression—the remaining population opportunity variable set reflects to a certain degree the shared impact of both population opportunity and employment-based opportunity. All eight of the dichotomous variables in the population opportunity set being highly significantly related to VMT, their coefficients *fulfill* the research objective—discovering the VMT impact of each level of proximity—and can therefore be used by government to score candidate SGAs according to the expected VMT benefit of their proximity level.

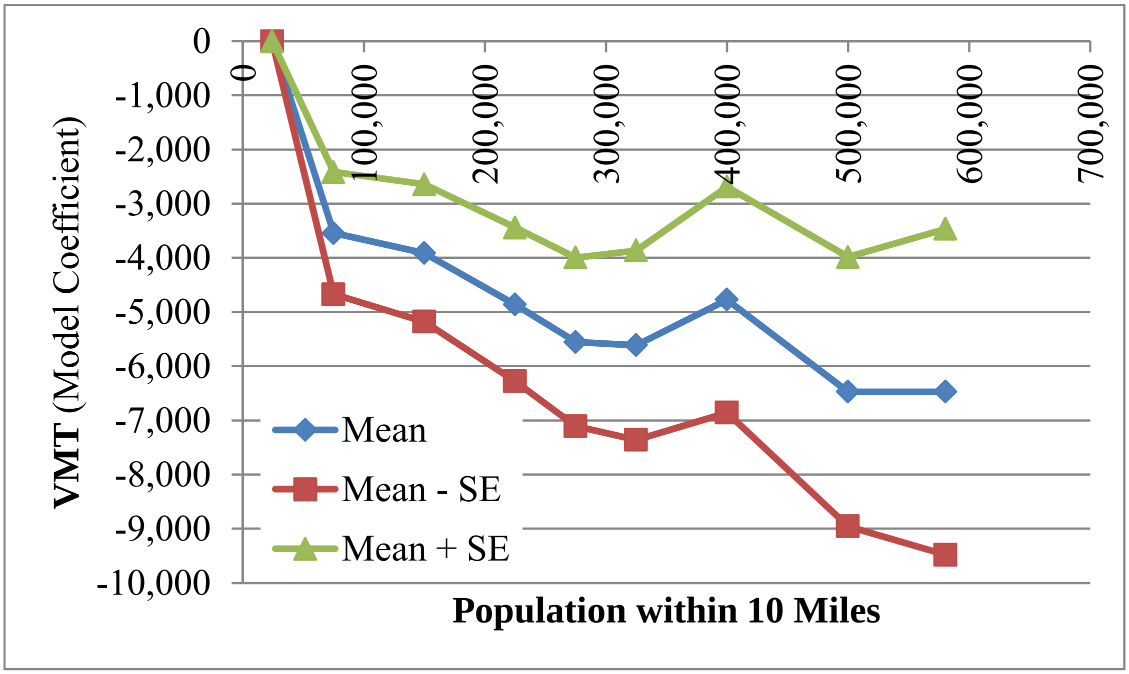


VMT-proximity curves.xlsx

FIGURE 16 VMT vs. Opportunity.

Although the above curve does not exhibit the theoretical flattening at lower proximity, the VMT curve does flatten at the right side as expected and discussed in "The Expected Shape of VMT-Proximity Curves and Key Hypothesis" section above. This flattening provides hope that a sweet spot may be located on the curve.

In preparation of testing the key hypothesis using the coefficients of the opportunity variables shown on the above curve, 1) the curve is re-plotted below showing standard errors (SE), and 2) the prevalence of the various proximity levels is explored in the table below.



VMT-proximity curves.xlsx

FIGURE 17 VMT vs. Opportunity.

TABLE 24 Prevalence of Opportunity Levels in Hampton Roads

Population within 10 Miles	**Count of Households in Hampton Roads, 2009**	**%**	**Household Count,** percentile range	
2009 Pop. within 10mi, 0k-50k	38,512	6%	0%	6%
2009 Pop. within 10mi, 50k-100k	47,715	8%	6%	14%
2009 Pop. within 10mi, 100k-200k	65,681	11%	14%	25%
2009 Pop. within 10mi, 200k-250k	70,962	12%	25%	37%
2009 Pop. within 10mi, 250k-300k	88,761	15%	37%	52%
2009 Pop. within 10mi, 300k-350k	82,081	14%	52%	65%
2009 Pop. within 10mi, 350k-450k	78,710	13%	65%	78%
2009 Pop. within 10mi, 450k-550k	84,021	14%	78%	92%
2009 Pop. within 10mi, 550k+	45,782	8%	92%	100%
	602,224	100%		

2009_TAZ_data.xlsx

The key hypothesis of this dissertation is:

> There exists a sweet spot on the VMT-proximity curve that has high VMT benefit *and* a proximity level acceptable to many households.

And the specific key hypothesis for testing is:

> The VMT benefit at 67% of maximum proximity is equal to or greater than 80% of the VMT benefit at maximum proximity.

Given that the average opportunity of the households in the highest opportunity level (550k+) is 580,571 persons within 10 miles, 67% of the 580,571 maximum proximity level is 388,983 persons within 10 miles. According to the above table, this 390k level is approximately the 70 percentile level of Hampton Roads households.

TABLE 25 Hypothesis Testing Worksheet based on VMT-Opportunity Curve

Specific Hypothesis:	The VMT benefit at 67% of max. proximity is >= 80% of the VMT benefit at max. proximity.		
Null Hypothesis:	The VMT benefit at 67% of max. proximity is < 80% of the VMT benefit at max. proximity.		
			source
Max. proximity (550k+ within 10 mi):	580,571	persons within 10 miles	VMT curve
	67%		
67% of max. proximity:	388,983	persons within 10 miles	
Mean VMT benefit @ 67% of max. prox.:	**4,772**	miles	Regression Table
Mean VMT benefit @ max. prox.:	6,473	miles	Regression Table
	80%		
80% of mean VMT benefit @ max. prox.:	**4,337**	miles	
Therefore, mean VMT benefit at 67% of max. prox. is higher than 80% of mean VMT benefit at max. prox.			
Testing this result considering the standard errors (SE) of the two benefits being compared:			
t-test requirements:	"two normally distributed but independent populations, σ is unknown" (*10*)		
	The two populations are mostly independent of each other and σ is unknown.		
Difference in the two benefits:	435	miles	
SE of VMT benefit @ 67% of max. prox.:	2,084	miles	Regression Table
SE of VMT benefit @ max. prox.:	3,009	miles	Regression Table
	80%		
SE of 80% of VMT benefit @ max. prox.:	2,407	miles	
Calculated t:	0.14	(calculated via formula for t for comparing two means)	
	vs.		
Critical t value:	1.28	(for α=0.10 and df >1,000)	
Therefore, the null hypothesis is not rejected.			

tables.xlsx

Based on the above hypothesis testing worksheet for the VMT-opportunity curve:

It is *likely* that the VMT benefit at 67% of maximum proximity is higher than 80% of the VMT benefit at maximum proximity, but—because the null hypothesis was not rejected—it is not *certain* that the VMT benefit at 67% of maximum proximity is higher than 80% of the VMT benefit at maximum proximity.

Only similar research on other metros would reveal whether these Hampton Roads findings are transferrable.

Given that the proximity level tested in the hypothesis test falls at the 70% of households (as stated above), i.e. a moderate level, it was not necessary to examine an additional, more moderate point on the curve as done in Effort #1 above.

VMT vs. Centrality

Regression Analysis As in the VMT-vs.-opportunity analysis above (where the VMT impact of each level of opportunity was determined, to determine the VMT impact of each level of centrality, OLS regression was used. The OLS regression was run using a set of binary variables based on the “Distance to Center” variable discussed in the Data Preparation section above and the control variables from the above VMT-opportunity regression, except for the statistically insignificant accessibility variable “2009 Pop. within 20min.” The results are shown below.

TABLE 26 VMT-Centrality OLS Regression Results

Source	SS	df	MS
Model	2.5E+11	22	1.1E+10
Residual	3.5E+11	2418	1.4E+08
Total	6.0E+11	2440	2.5E+08

Number of obs	2,441
F(32, 8926)	78.84
Prob > F	0.000
R-squared	0.418
Adj R-squared	0.412
Root MSE	12,009

DV: ANNMILES	Coef.	Std. Err.	t	P> |t|	Signif*	95% Conf. Interval	
Control Variables							
Household Family Income							
Basis: HHFAMINC <$20k							
HHFAMINC missing	1,577	1,201	1.31	0.189	--	-778	3,931
HHFAMINC $20,000-$39,999	1,492	974	1.53	0.126	--	-418	3,402
HHFAMINC $40,000-$59,999	2,899	1,016	2.85	0.004	√√	907	4,891
HHFAMINC $60,000-$99,999	4,784	1,029	4.65	0.000	√√	2,767	6,801
HHFAMINC $100,000+	8,272	1,110	7.45	0.000	√√	6,095	10,448
Home Owned	2,995	801	3.74	0.000	√√	1,424	4,566
All Household Members (Age 5+)							
Male Workers (Age 16+)	10,984	566	19.42	0.000	√√	9,875	12,094
Female Workers (Age 16+)	8,452	565	14.95	0.000	√√	7,343	9,561
Male Non-Workers (Age 16+)	5,970	590	10.12	0.000	√√	4,814	7,127
Female Non-Workers (Age 16+)	5,430	590	9.21	0.000	√√	4,274	6,587
Persons Age 5 thru 15	514	386	1.33	0.183	--	-243	1,271
Persons 16+ Having MEDCOND	-2,101	584	-3.60	0.000	√√	-3,246	-956
Persons 16+ Never Used Internet in Past Mo.	-2,204	464	-4.75	0.000	√√	-3,114	-1,295
Policy Variables							
Centrality							
Basis: Distance to Center, <4mi							
Distance to Center, 4-6mi	1,559	1,080	1.44	0.149	--	-560	3,677
Distance to Center, 6-8mi	2,118	1,050	2.02	0.044	√√	60	4,176
Distance to Center, 8-10mi	2,109	1,036	2.04	0.042	√√	78	4,140
Distance to Center, 10-12mi	2,630	1,062	2.48	0.013	√√	549	4,712
Distance to Center, 12-14mi	2,491	1,239	2.01	0.044	√√	62	4,919
Distance to Center, 14-18mi	4,704	1,065	4.41	0.000	√√	2,614	6,793
Distance to Center, 18-25mi	5,000	1,228	4.07	0.000	√√	2,593	7,407
Distance to Center, 25-30mi	2,790	1,216	2.29	0.022	√√	405	5,174
Distance to Center, 30+ mi	8,645	1,119	7.73	0.000	√√	6,451	10,839
Constant	-4,728	1,276	-3.71	0.000	√√	-7,230	-2,226

* "√": Significant at the 0.10 level; "√√": Significant at the 0.05 level

Tables.xlsx

With an R-squared value exceeding 0.4, the above regression shows an excellent statistical relationship between the theory-based set of independent variables and annual household VMT. Concerning the control variables, most are statistically significant and have coefficients with logical size and sign. Concerning the policy variables, nine out of ten of the binary variables are statistically significant at the 0.05 level. Prior to discussing the VMT-centrality regression results, the threats to its validity will be addressed.

Threats to Validity Threats to the validity of the model resulting from the above regression process were checked by addressing the following topics:

- Logical coefficient signs and values
- Influence points
- Normality
- Homoscedasticity
- Linearity
- Independence of error terms
- Model fit
- Self-selection

Logical Coefficient Signs and Values The control variables in this regression being the same as the control variables in the above VMT-opportunity regression (except for having dropped the accessibility variable), and coefficients being very similar between the two models, the coefficient signs and values of this regression are again logical.

Influence Points The maximum value of the seven scalar variables being 5, there are no outliers, eliminating the concern over undue influence from stray low or high values.

Normality As in the VMT-opportunity regression above, the sample size of the VMT-centrality model (2,441) exceeding 200 observations, the issue of normality was considered not to be problematic.

Homoscedasticity and Linearity The centrality set of policy variables being dichotomous, homoscedasticity and linearity are not a concern. In fact, the theorized non-linearity of the relationship between proximity and VMT was the purpose of creating the set of dichotomous centrality variables.

Independence of Error Terms As in the VMT-opportunity regression above, given the use of *annual* VMT for the dependent variable of the VMT-centrality regression, sequencing (i.e. the date each survey was taken) is not a concern.

Model Fit In addition to the fact that most of the variables in the models (including nine out of ten of the policy variables) are significantly related to annual VMT (Type I error rate < 0.05), the Adjusted R-squared value is 0.41, demonstrating an excellent model fit.

Travel Mode Biases ("self-selection") Travel mode biases were addressed in this effort in the Brownstone (12) manner discussed in the Preparation section above, i.e. by including several key socio-economic variables in the model.

Overall Assessment of the Model Given the satisfactory survey of the threats to model validity, it appears that the model is reliable for use in investigating VMT impact by centrality level.

Useful Regression Results and Hypothesis Testing The coefficients of the control variables being very similar to those of the earlier VMT-opportunity regression (and the implications of these control variable coefficients having been discussed under that regression above), only the policy variable results are discussed here. Eight of the nine dichotomous variables in the centrality set being highly significantly related to VMT (and the ninth variable being significant at the 0.10 level), their coefficients can be used by government to score candidate SGAs according to the expected VMT benefit of their proximity level.

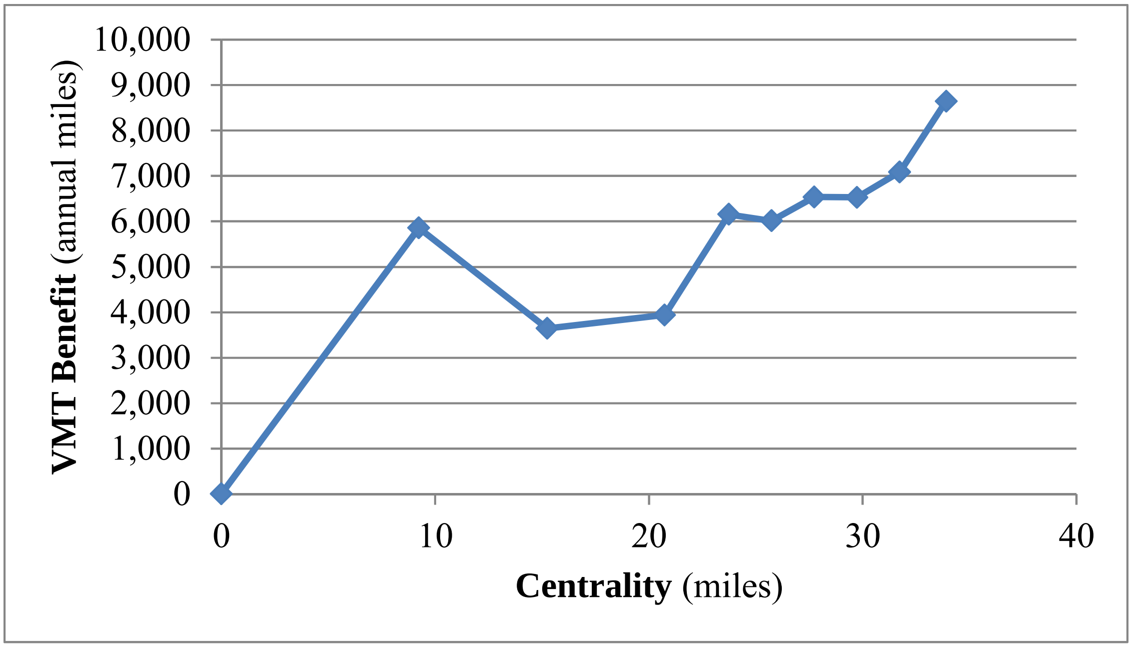
The set of policy variables in the subject regression being based on distance-to-center, the above coefficients indicate the VMT *dis*benefit as compared to the highest level of centrality, <4 miles distance-to-center. In order to couch the result in terms of VMT *benefit*, as done for all the other regressions in this dissertation, the following table was prepared by subtracting the model coefficients from the VMT disbenefit at the lowest level of centrality (30+ miles distance-to-center, 8,645 miles VMT disbenefit).

TABLE 27 VMT Benefit at Each Level of Centrality (as compared to lowest level)

Distance to Center	**Centrality** (distance of most distant level - distance of level), miles	**Model Coef.,** miles	**VMT Benefit,** miles
Basis: Distance to Center, <4mi (avg. 2.83)	33.91	n.a.	8,645
Distance to Center, 4-6mi	30.74-32.74	1,559	7,086
Distance to Center, 6-8mi	28.74-30.74	2,118	6,527
Distance to Center, 8-10mi	26.74-28.74	2,109	6,536
Distance to Center, 10-12mi	24.74-26.74	2,630	6,015
Distance to Center, 12-14mi	22.74-24.74	2,491	6,154
Distance to Center, 14-18mi	18.74-22.74	4,704	3,941
Distance to Center, 18-25mi	11.74-18.74	5,000	3,645
Distance to Center, 25-30mi	6.74-11.74	2,790	5,855
Distance to Center, 30+ mi (avg. 36.74)	0	8,645	0

Tables.xlsx

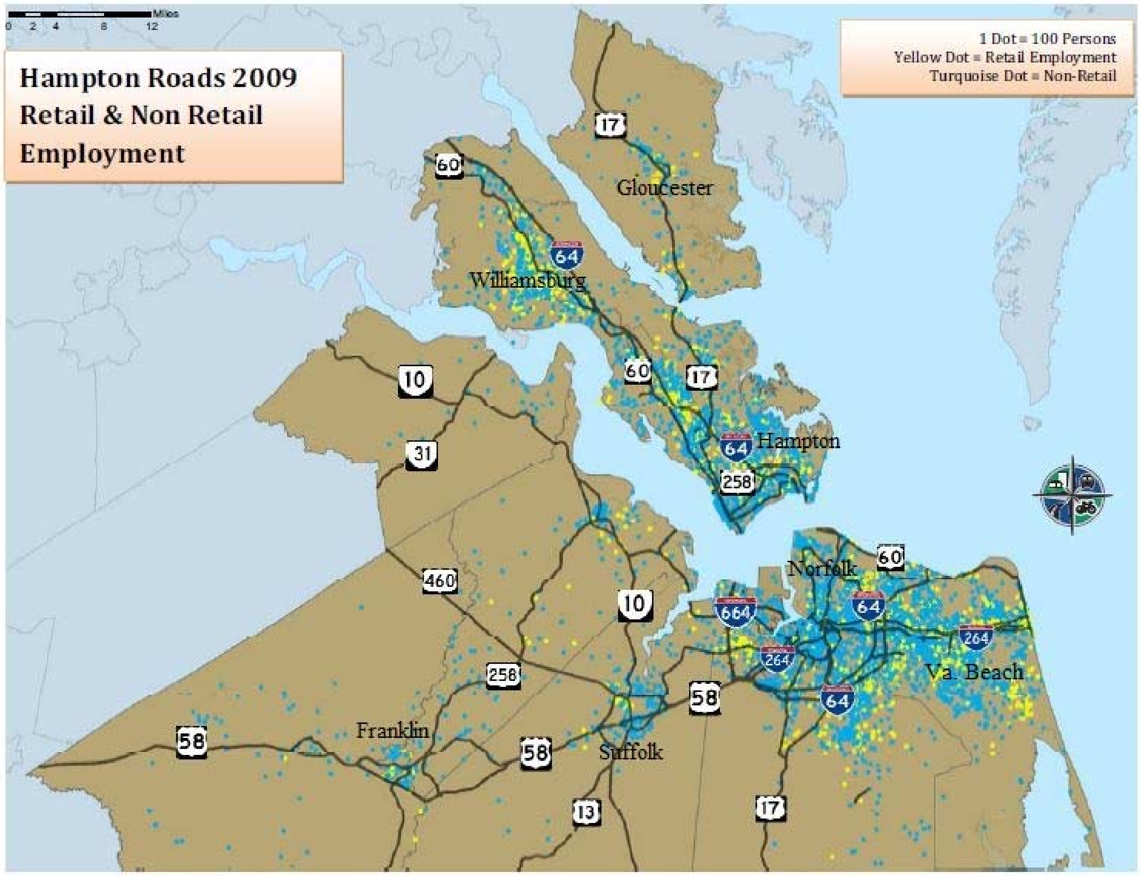
A curve based on this table can be found below.



VMT-centrality curves.xlsx

FIGURE 18 VMT vs. Centrality.

The VMT curve exhibits neither the flattening at lower centrality nor the flattening at higher proximity expected and discussed in "The Expected Shape of VMT-Centrality Curves and Secondary Hypothesis" section above. Except for the data point for the 6.74-11.74 miles centrality level (25-30 miles distance-to-center) discussed below, the curve is fairly linear. This linearity provides little hope that a sweet spot may be located on the curve.

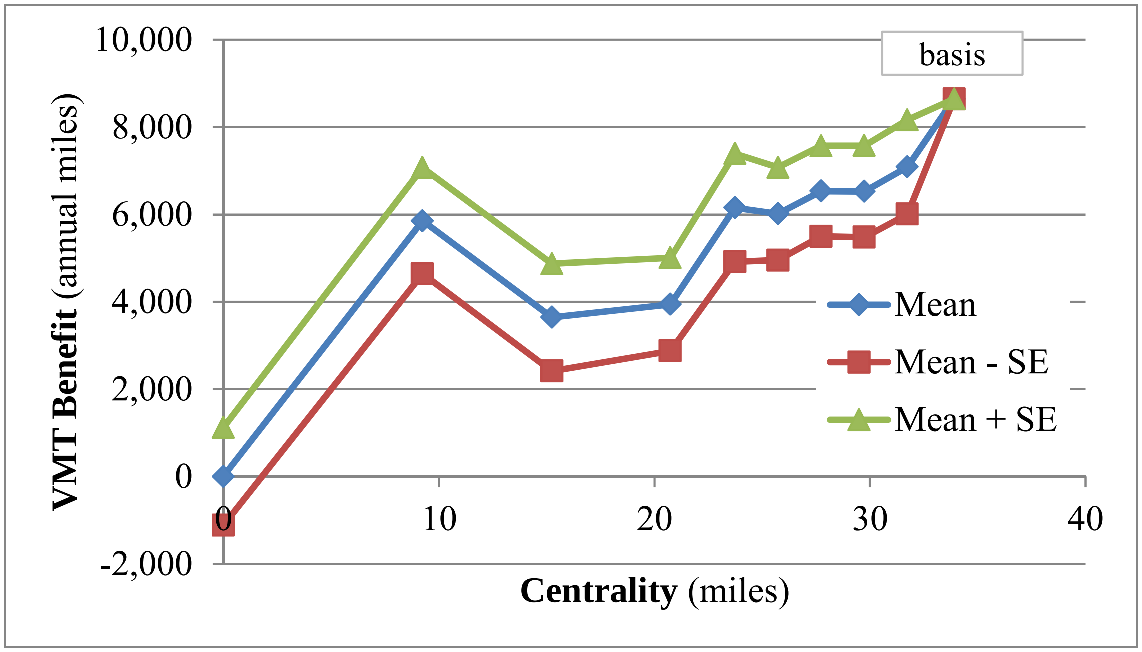


2009 HR Employment.jpg source: HRTPO (*34*), modified by adding place labels

FIGURE 19 Hampton Roads Employment Locations

An examination of the 170 surveyed households situated 25-30 miles from center revealed that 149 (or 88%) are located near Williamsburg and downtown Suffolk, two significant employment sub-centers, as shown on the above map. It is assumed that this proximity to these sub-centers explains the anomalous data point for 25-30 miles from center shown on the above graph.

In preparation of testing the secondary hypothesis using the coefficients from the VMT-vs.-centrality regression and curve, 1) the curve is re-plotted below with standard errors (SE), and 2) the prevalence of the various centrality levels in Hampton Roads is provided in the table below.



VMT-centrality curves.xlsx

FIGURE 20 VMT vs. Centrality.

TABLE 28 Prevalence of Population Centrality Levels in Hampton Roads

Population within 10 Miles	**Centrality** (distance of most distant level - distance of level), miles	**Count of Households in Hampton Roads, 2009**	**%**	**Household Count,** percentile range	
2009 Pop. within 10mi, 0k-50k	33.91	38,512	6%	0%	6%
2009 Pop. within 10mi, 50k-100k	30.74-32.74	47,715	8%	6%	14%
2009 Pop. within 10mi, 100k-200k	28.74-30.74	65,681	11%	14%	25%
2009 Pop. within 10mi, 200k-250k	26.74-28.74	70,962	12%	25%	37%
2009 Pop. within 10mi, 250k-300k	24.74-26.74	88,761	15%	37%	52%
2009 Pop. within 10mi, 300k-350k	22.74-24.74	82,081	14%	52%	65%
2009 Pop. within 10mi, 350k-450k	18.74-22.74	78,710	13%	65%	78%
2009 Pop. within 10mi, 450k-550k	11.74-18.74	84,021	14%	78%	92%
2009 Pop. within 10mi, 550k+	6.74-11.74	45,782	8%	92%	100%
	0	602,224	100%		

2009_TAZ_data.xlsx

The secondary hypothesis of this dissertation is:

> There exists a sweet spot on the VMT-centrality curve that has high VMT benefit *and* a centrality level acceptable to many households.

And the specific secondary hypothesis for testing is:

> The VMT benefit at 67% of maximum centrality is equal to or greater than 80% of the VMT benefit at maximum centrality.

Given that the average centrality of the households in the highest centrality level (0-4 miles distance from center) is 33.91 miles (as shown on the table above), 67% of the 33.91 level is 22.72 miles, which falls in the 18.74-22.74 centrality level. According to the above table, this 22.72 level is approximately the 65 percentile level of Hampton Roads households.

TABLE 29 Hypothesis Testing Worksheet based on VMT-Centrality Curve

Specific Hypothesis:	The VMT benefit at 67% of max. centrality is >= 80% of the VMT benefit at max. centrality.		
Null Hypothesis:	The VMT benefit at 67% of max. centrality is < 80% of the VMT benefit at max. centrality.		
			source
Max. centrality (<4mi distance to center):	33.91	miles	VMT curve
	67%		
67% of max. centrality:	22.72	miles	(centrality 18.74-22.74 mi.; distance to center: 14-18 mi.)
Mean VMT benefit @ 67% of max. cent.:	**3,941**	miles	Regression Table
Mean VMT benefit @ max. cent.:	8,645	miles	Regression Table
	80%		
80% of mean VMT benefit @ max. cent.:	**5,792**	miles	

Therefore, mean VMT benefit at 67% of max. centrality is less than 80% of mean VMT benefit at max. centrality.

Given the above result, there is no need to conduct a t-test to test the null hypothesis.

Therefore, the null hypothesis is not rejected.

tables.xlsx

Based on the above hypothesis testing worksheet for the VMT-centrality curve—and as feared given the linearity of the VMT-centrality curve:

It is *very unlikely* that the VMT benefit at 67% of maximum centrality is higher than 80% of the VMT benefit at maximum centrality.

Given that centrality is a proxy for proximity—as opposed to a true measure of proximity—this finding concerning the secondary hypothesis does not negate the earlier positive findings in this dissertation concerning the key hypothesis that there exists a sweet spot on the VMT-proximity curve that has high VMT benefit and a proximity level acceptable to many households.

Given the varying sizes of U.S. metros, it is doubtful that the slope and intercept of the Hampton Roads VMT-centrality curve are transferrable to other metros. It may be, however, that the centrality-VMT relationship is roughly linear in many metros.

CHAPTER V

CONCLUSION

Fulfillment of Research Objective- The VMT Impact of Each Proximity Level

The research objective of this dissertation is:

to discover the VMT impact of each level of proximity.

Given that the empirical research in Efforts #1 and #3 above discovered the VMT impact of each level of proximity, the research objective of this dissertation was *fulfilled*.

The coefficients of the two sets of density variables (population and employment) in Effort #1 reveal the VMT benefit at each level of census tract density. For example, the regression revealed that the 50-99 per-square-mile (census tract) level of employment density is associated with a VMT benefit of 400 annual miles (as compared to the lowest employment level), and the 100-499 per-square-mile (census tract) level of population density is associated with a VMT benefit of 1,700 annual miles (as compared to the lowest population level).

The coefficients of the sets of policy variables in the two analyses in Effort #3 reveal—using the first analysis—the relationship between opportunity (in this case, population within 10 miles) and VMT at each level of opportunity, and—using the second analysis—the relationship between centrality and VMT at each level of centrality. For example, the opportunity regression revealed that the 50,000-100,000 persons-within-10-miles level of opportunity is associated with a VMT benefit of 3,500 annual miles (as compared to the lowest opportunity level). And the centrality regression revealed that the 4-6 miles-from-center level of centrality is associated with a VMT benefit of 7,000 annual miles (as compared to the lowest centrality level).

Fulfillment of Dissertation Purpose and Application of Dissertation Models

The purpose of this dissertation is:

> to discover the VMT impact of each level of proximity in order to help government identify key locations for housing development, and thereby lower VMT and reduce dependence on foreign oil.

Given that this dissertation's discovery of the VMT impact of each level of proximity can be applied—using the technique described below—to help government identify key locations for housing development, the purpose of this dissertation has been *fulfilled*.

Application of the Dissertation Models

Governments can use the policy variable coefficients from the final models in Efforts #1 and #3—via the VMT Benefit Technique described below—to accurately determine the VMT benefit of a given location, and use the VMT benefits of a set of candidate areas to select key locations for development.

The VMT Benefit Technique The process of determining the VMT benefit of a given location, known herein as the "VMT Benefit Technique", is described as follows. Governments around the U.S. can use the coefficients from the policy variables in Effort #1 to accurately determine the VMT benefit of any U.S. location, and governments in Hampton Roads can use the coefficients from the policy variables in the models of Effort #3 to accurately determine the VMT benefit of any Hampton Roads location.

First, governments around the U.S. can apply the policy variable coefficients from the final regressions in Effort #1 to calculate the VMT benefit of any location in America. Simply:

- determine the population and employment densities (census tract level) of the subject location,
- look up the coefficients for those densities in the VMT-Density regression results table above, and
- add the population coefficient and the employment coefficient together to calculate the VMT benefit.

For example, a location with 6,432 persons per square mile and 1,233 employed persons per square mile is at the 4,000-9,999 population density level and the 1,000-1,999 employment density level, and has therefore a VMT benefit of 9,000 annual miles (4,549 + 4,320 = 8,869) vs. a location with the lowest density levels.

Likewise, local governments in Hampton Roads can apply the policy variable coefficients from the regressions in Effort #3 to calculate the VMT benefit of any location in Hampton Roads. And they can do so using either the opportunity analysis or the centrality analysis. Using the opportunity analysis, simply:

- determine the TAZ of the subject location (by examining the TAZ document available on the website of the Hampton Roads Transportation Planning Organization [HRTPO]),
- look up the amount of population within 10 miles of that TAZ (using a table developed for this dissertation), and
- look up the coefficient for that opportunity level in the VMT-Opportunity regression results table above, the coefficient being the VMT benefit of the subject location.

For example, a location in TAZ 1 has 514,503 persons within 10 miles. Looking up the coefficient for the 450,000-550,000 opportunity level, indicates that the subject location has a VMT benefit of 6,500 annual miles (coefficient 6,470). Alternately, using the centrality analysis in Effort #3, simply:

- determine the TAZ of the subject location (by examining the TAZ document available on the website of the HRTPO),
- calculate the distance (perhaps using Google maps) from that TAZ to the appropriate center (the Ballentine-264 interchange on the Southside; the Peninsula Town Center on the Peninsula), and

- look up the VMT disbenefit for that centrality level in the VMT-Centrality regression results table above, or—more directly—look up the VMT benefit for that centrality level in the “VMT Benefit at Each Level of Centrality” table above.

For example, for the same location above (in TAZ 1), TAZ 1 is 2.79 miles from the Ballentine-264 interchange. Looking up the coefficient for the <4 mile centrality level in the “VMT Benefit at Each Level of Centrality” table above indicates that the subject location has a VMT benefit of 8,645 annual miles. Note that, for TAZ 1, the two models—the VMT-opportunity model and the VMT-proximity model—appropriately give similar results.

Using VMT Benefit Technique to Consider VMT in Choosing SGAs Governments around the U.S. and in Hampton Roads can apply this VMT Benefit Technique to identify locations in which they would prefer development occur, e.g. strategic growth areas (SGAs). They can locate a set of candidate SGAs, use the coefficients from any of the Effort #1 and Effort #3 models to calculate the VMT benefit of each candidate area, and use those results as one input to the process of choosing final SGAs, i.e. “to identify key locations for development” as in the dissertation title. Whereas government currently considers many non-VMT factors when identifying SGAs—e.g. availability of land for development or redevelopment, existing supportive infrastructure, etc.—by using the coefficients of the final models in Efforts #1 and #3 to estimate the expected VMT impacts of the proximities of the locations of the candidate SGAs, it can add VMT reduction as a factor in the process of identifying key locations for development.

Key Hypothesis and Implication of the Coefficients in the Dissertation Models

The key hypothesis of this dissertation is:

> There exists a sweet spot on the VMT-proximity curve that has high VMT benefit and a proximity level acceptable to many households.

Given the hypothesis tests conducted using the results of the preceding empirical analyses, the key *null* hypothesis of this dissertation is *not rejected*, i.e. it is *not certain* that the sweet spot exists. However, the mean coefficients of each VMT-proximity analysis in this dissertation indicate that it is *likely* that there are high-VMT-benefit proximity levels acceptable to many households, i.e. that the sweet spot exists. The overall implication of this is that representative governments in the U.S. who promote housing development at these moderate levels of proximity will not only lower average VMT in the short term, they will not be punished politically for doing so, and therefore may be successful in thereby lowering average VMT in the long term.

Key Implications, Primary Contribution to the Literature, and Long-term Value

The key implications, primary contribution to the literature, and long-term value of this dissertation is that:

> a) it provides encouragement to governments hoping to lower average VMT, and
>
> b) it provides an accurate method of calculating VMT for choosing SGAs with which to actually lower average VMT.

Future Research Directions

The findings of Efforts #1 and #2, being based on a nationwide dataset, are applicable nationwide, but this dataset lacked details available on the local level, such as a rigorous measurement of proximity and a measure of the availability of public transportation. The findings of Effort #3, being based on a local dataset, are only applicable locally. Therefore, future research of several dissimilar metros (including some having significant usage of public transit) that includes a rigorous measurement of proximity (e.g. the opportunity and centrality used in Effort #3) and a measure of the availability of public transportation could provide widely applicable results.

Final Conclusion

Given that many Americans would dislike living in high proximity locations known for having a low VMT signature, and that American government has a representative form, the lack of VMT benefit information by individual proximity level in the literature made the application of the "higher-proximity equals lower-VMT" message of the existing literature difficult to apply to date. Fortunately, this dissertation discovered the VMT impact of each level of proximity.

Governments can apply the discovered VMT impact of each level of proximity—via a described "VMT Benefit Technique"—to accurately determine the VMT benefit of a given location, and use the VMT benefits of a set of candidate areas to select key locations for development.

In addition, the discovered VMT impact of each level of proximity informs the key hypothesis of this dissertation that there exists a sweet spot on the VMT-proximity curve that has high VMT benefit and a proximity level acceptable to many households. Although the hypothesis tests indicate that it is *not certain* that the sweet spot exists, the mean coefficients of the models indicate that it is *likely* that the sweet spot exists, i.e. that there are high-VMT-benefit proximity levels acceptable to many households. The overall implication of this is that representative governments in the U.S. who promote housing development at these moderate levels of proximity will not only lower average VMT in the short term, they will not be punished politically for doing so, and therefore may be successful in thereby lowering average VMT in the long term.

In summary, the dissertation provides encouragement to governments hoping to lower average VMT and an accurate method of calculating VMT for choosing SGAs with which to actually lower average VMT. It is hoped that this combination will help U.S. governments become independent of foreign oil.

REFERENCES

(1) *Driving and the Built Environment*, Special Report 298, Transportation Research Board, Washington, DC, 2009, pp. 55-90.

(2) Badoe, Daniel A., and Eric J. Miller. Transportation-land use interaction: empirical findings in North America, and their implications for modeling. *Transportation Research*, Part D, Vol. 5, No. 4, 2000, pp. 235-263.

(3) Cervero, Robert and Michael Duncan. Which Reduces Vehicle Travel More: Jobs-Housing Balance or Retail-Housing Mixing? *Journal of the American Planning Association*, Vol. 72, No. 4, Autumn 2006, pp. 475-490.

(4) Case, Robert. *Improving the Mobility of Non-Drivers Using Proximity to Destinations and Bus Routes.* Technical Document #T07-06. Hampton Roads Planning District Commission, June 2007, p. 19.

(5) Khattak, Asad J., Virginie J. Amerlynck, and Roberto G. Quercia. Are Travel Times and Distances to Work Greater for Residents of Poor Urban Neighborhoods? In *Transportation Research Record: Journal of the Transportation Research Board*, No. 1718, Transportation Research Board of the National Academies, Washington, D.C., 2000, pp. 73-82.

(6) Cervero, Robert and Jin Murakami. Effects of built environments on vehicle miles traveled: evidence from 370 US urbanized areas. *Environment and Planning A*, Vol. 42, No. 2, 2010, pp. 400-418.

(7) Pushkarev, Boris S., and Jeffrey M. Zupan. *Public Transportation and Land Use Policy.* Indiana University Press, Bloomington & London, 1977, pp. 29-35.

(8) *Measuring What Matters: Access to Destinations*, Access to Destinations Study Research Summary No. 2, Center for Transportation Studies, University of Minnesota, August 2010, p. 3-5.

(9) Miller, David R. and Ken Hodges. A Population Density Approach to Incorporating an Urban-Rural Dimension into Small Area Lifestyle Clusters. Presented at the Annual Meeting of the Population Association of America, Miami, May 4-7, 1994.

(10) Voelker, David H., Peter Z. Orton, and Scott V. Adams. *CliffsQuickReview Statistics*. Wiley Publishing, Inc., New York, 2001, p. 84.

(11) Hair, Joseph F., William C. Black, Barry J. Babin, Rolph E. Anderson, and Ronald L. Tatham. *Multivariate Data Analysis*. Pearson Prentice Hall, Upper Saddle River, NJ, 2006.

(12) Brownstone, David. 2008. "Key Relationships Between the Built Environment and VMT." A commissioned paper for *Driving and the Built Environment*. Special Report 298. Transportation Research Board. 2009, p. 2.

(13) Schimek, Paul. Household Motor Vehicle Ownership and Use: How Much Does Residential Density Matter?. In *Transportation Research Record*, No. 1552, Transportation Research Board of the National Academies, Washington, DC, 1996, pp. 120-130.

(14) Ortuzar, J. de D., and Luis G. Willumsen. *Modelling Transport*. John Wiley & Sons, Ltd., Chichester, England, 2001.

(15) *A Transportation Research Program for Mitigating and Adapting to Climate Change and Conserving Energy*. Transportation Research Board Special Research Report 299. Transportation Research Board, Washington, DC, 2009.

(16) Dunphy, Robert T., and Kimberly Fisher. Transportation, Congestion, and Density: New Insights. In *Transportation Research Record*, No. 1552, Transportation Research Board of the National Academies, Washington, DC, 2003, pp. 89-96.

(17) *2009 NHTS User's Notes*. Federal Highway Administration. http://nhts.ornl.gov/2009/pub/UserNotes.pdf. Accessed Jan. 26, 2013, no pg. numbers.

(18) Bento, Antonio M., Maureen L. Cropper, Ahmed Mushfiq Mobarak, and Katja Vinha. 2005. "The Effects of Urban Spatial Structure on Travel Demand in the United States." *Review of Economics & Statistics* 87, no. 3: 466-478. *Business Source Complete*, EBSCO*host* (accessed September 1, 2011).

(19) Kockelman, Kara Maria. Travel Behavior as a Function of Accessibility, Land Use Mixing, and Land Use Balance. In *Transportation Research Record*, No. 1607, Transportation Research Board of the National Academies, Washington, DC, 1997, pp. 116-125.

(20) Walls, M., W. Harrington, and A. Krupnick. Population Density, Transit Availability, and Vehicle Travel: Results from a Nested Logit Model of Vehicle Ownership and Use. Draft manuscript, Resources for the Future, Washington, DC, Jan. 2002.

(21) Golob, Thomas F., and David Brownstone. *The Impact of Residential Density on Vehicle Usage and Energy Consumption*. Publication UC-ITS-WP-05-01. Institute of Transportation Studies, Irvine, CA, Feb. 2005.

(22) Chen, Yali, et al. Development of Opportunity-Based Accessibility Indicators. Submitted for 90th Annual Meeting of the Transportation Research Board, Washington, DC, Jan. 2011, p. 3.

(23) Ross, Catherine L, and Anne E. Dunning. *Land Use Transportation Interaction: An Examination of the 1995 NPTS Data*. Georgia Institute of Technology. Prepared for FHWA, Washington, DC, 1997.

(24) Shay, Elizabeth. *Homes, autos, and travel: Household decision chains*. Dissertation, University of North Carolina, Chapel Hill, 2007.

(25) Salon, Deborah, Marlon G. Boarnet, Susan Handy, Steven Spears, and Gil Tal. How do local actions affect VMT? A critical review of the empirical evidence. Presented at 91st Annual Meeting of the Transportation Research Board, Washington, DC, 2012.

(26) Yoon, Seo Youn, Thomas F. Golob, and Konstadinos G. Goulias. A California Statewide Exploratory Analysis Correlating Land Use Density, Infrastructure Supply and Travel Behavior. CD-ROM from 88th Annual Meeting of the Transportation Research Board, Washington, DC, 2009.

(27) Handy, Susan. Regional Versus Local Accessibility: Implications for Nonwork Travel. In Transportation Research Record, No. 1400, Transportation Research Board of the National Academies, Washington, DC, 1993, pp. 58-66.

(28) Pickard, Andrew. *A Review of 2000 Census Commute Data for Hampton Roads.* Technical Document #T05-03. Hampton Roads Planning District Commission, Nov. 2005, p. 69.

(29) Boarnet, Marlon G., Douglas Houston, Gavin Ferguson, and Steven Spears. Land Use and Vehicle Miles of Travel in the Climate Change Debate: Getting Smarter than Your Average Bear. Presented at the 5th annual Land Policy Conference of the Lincoln Institute of Land Policy, Cambridge, Massachusetts, 2010.

(30) Cao, Xinyu, Zhiyi Xu, and Yingling Fan. Exploring the connections among residential location, self-selection, and driving: A propensity score matching approach. CD-ROM from 89th Annual Meeting of the Transportation Research Board, Washington, DC, 2010.

(31) Zhou, Bin (Brenda), and Kara M. Kockelman. Self-Selection in Home Choice: Use of Treatment Effects in Evaluating Relationship Between Built Environment and Travel Behavior. In Transportation Research Record, No. 2077, Transportation Research Board of the National Academies, Washington, DC, 2008, pp. 54-61.

(32) Krizek, Kevin J. Residential Location and Changes in Urban Travel: Does Neighborhood-Scale Urban Form Matter? Journal of the American Planning Association, Vol. 69, No. 3, Summer 2003, pp. 265-281.

(33) Shay, Elizabeth, and Asad J. Khattak. Toward Sustainable Transport: Conventional and Disruptive Approaches in the U.S. Context. In International Journal of Sustainable Transportation, Vol. 4, 2010, pp. 1-27, draft copy.

(34) Pickard, Andrew. *Hampton Roads 2009 Socioeconomic Data by TAZ.* Technical Document #T10-08. Hampton Roads Transportation Planning Organization, Nov. 2010, p. 2.

(35) Duany, Andrés. A New Theory of Urbanism. Scientific American, Vol. 283, No. 6, Dec. 2000, pp. 251-260.

(36) Banerjee, Neela. Obama seeks to reduce America's oil imports by one-third. McClatchy-Tribune Information Services. www.azdailysun.com/news/national. Accessed April 12, 2011.

(37) Morin, Jim. Editorial cartoon. The Miami Herald. www.democraticunderground.com/10165415. Accessed August 13, 2012.

(38) Kahn, Chris. As We Pay More for Gas, Fuel Becomes Top U.S. Export. *The Virginian-Pilot*, Dec. 31, 2011, p. A1.

(39) Banerjee, Neela. U.S. oil output highest since 2003, report finds. *Daily Press*. Mar. 13, 2012, p. 13 (News).

(40) Dictionary. Merriam-Webster. http://www.merriam-webster.com/dictionary/travel. Accessed August 22, 2012 and February 19, 2013.

(41) Schneider, Morton. Gravity Models and Trip Distribution Theory. *Papers in Regional Science*, Vol. 5, Issue 1, 1959, pp. 51-56.

(42) U.S. Department of Transportation, Federal Highway Administration. *2009 National Household Travel Survey.* URL: http://nhts.ornl.gov.

(43) U.S. Department of Transportation, Federal Highway Administration. *Developing a Best Estimate of Annual Vehicle Mileage for 2009 NHTS Vehicles.* URL: http://nhts.ornl.gov/2009/pub/BESTMILE.pdf, no page numbers.

VITA

Robert B. Case, P.E.

Department of Civil and Environmental Engineering
Kaufman Hall 135
College of Engineering and Technology
Old Dominion University
Norfolk, VA 23529-0241

Robert B. Case is a Principal Transportation Engineer with the Hampton Roads Transportation Planning Organization (HRTPO), where he leads a team of engineers who staff the Hampton Roads Transportation Operations subcommittee and perform congestion analyses, military studies, and freight studies. In addition, through the development and application of mathematical models, Mr. Case has performed work in areas such as long-range transportation planning, non-driver mobility, and hurricane evacuation. Mr. Case is a licensed Professional Engineer in Virginia, a registered Professional Traffic Operations Engineer, and an ASCE Fellow. He has been an adjunct instructor at Old Dominion University where he has helped teach transportation planning and currently helps teach a class for engineers preparing to take the PE exam. He received a B.A. in History from the University of Virginia, a B.S. in Civil Engineering from Old Dominion University, and an M.S. in Civil Engineering from Georgia Tech. His articles were published in Transportation Research Record No. 1502 (1995) and No. 2242 (2011). He received the John T. Hanna Award from Concerned Citizens Advocating Traffic Safety in 1996 and the Innovative Intermodal Solutions for Urban Transportation Award from ITE in 2006.