Modeling Household Vehicle and Transportation Choice and Usage

Part A: Factors Related to Voluntary Choice of Low Vehicle Ownership and Usage

Project Report Prepared for

State of California Air Resources Board
Research Division
PO Box 2815
Sacramento CA 95812

Prepared by:

Prof. Patricia L. Mokhtarian, co-Principal Investigator

David Van Herick

Cheng Zhuo

and

Dr. Giovanni Circella

Institute of Transportation Studies
University of California, Davis
One Shields Avenue
Davis, CA 95616

April 2017

DISCLAIMER

The statements and conclusions in this Report are those of the contractor and not necessarily those of the California Air Resources Board. The mention of commercial products, their source, or their use in connection with material reported herein is not to be construed as actual or implied endorsement of such products.

ACKNOWLEDGEMENTS

This project was funded by the Research Division of the State of California Air Resources Board. The contents of this report reflect the views of the authors, and not necessarily those of the funding agency. We gratefully acknowledge the substantial assistance of Farzad Alemi in estimating some of the models associated with Task A.5. While we received feedback during this research from numerous individuals, the authors are responsible for the facts and the accuracy of the information presented, and any errors or omissions are the responsibility of the authors alone.

This Report, along with “Part B: Empirical Estimation of Household Vehicle Purchase and Usage Decisions”, was submitted in fulfillment of ARB Agreement Number 11-322, “Modeling Household Vehicle and Transportation Choice and Usage”, by the Regents of the University of California, Davis, under the sponsorship of the California Air Resources Board. The work on Part A was completed as of April 24, 2017.

Table of Contents

List of Figures

List of Tables

List of Acronyms

AAE	About as Expected
ARB	Air Resources Board
DV	Dummy Variable
EL	Equally Likely
GHG	Greenhouse Gas(es)
HDN	Higher-Density Neighborhood
HTE	Higher Than Expected
LDN	Lower-Density Neighborhood
LL	Log-Likelihood
LTE	Lower Than Expected
MSA	Metropolitan Statistical Area
NHTS	National Household Travel Survey
PMT	Person-Miles Traveled
PUMS	Public Use Microdata Sample
RTP	Regional Transportation Plan
SB	Senate Bill
SCS	Sustainable Community Strategy
US	United States
VMT	Vehicle-Miles Traveled
WLS	Weighted Least Squares
ZVO	Zero Vehicle Ownership

ABSTRACT

This study analyzes the 2009 National Household Travel Survey (NHTS) and attitudinal data pooled across four California studies, to investigate the impact of household demographics, individual attitudes, and residential location on vehicle ownership and usage decisions in California.

We classify each household as lower-than-expected (LTE), about-as-expected (AAE), or higher-than-expected (HTE) vehicle owning, based on the comparison of the actual vehicle ownership level with the expected value computed from a model that predicts vehicle ownership based on household size and composition. Households that do not own any vehicles are classified as zero-vehicle-owning (ZVO). We are especially interested in exploring the reasons for which a household would own fewer-than-expected or no vehicles, and have low vehicle miles traveled (VMT). We first estimate a set of models to explore the impacts of the most natural constraints that could explain low- or zero-vehicle-ownership status, and lower VMT, namely income and driving limitations. We then focus on the reasons for the voluntary choice of low/zero vehicle ownership through controlling for the impacts of personal attitudes. Finally, we analyze the role of residential location and land use traits in affecting household vehicle ownership and VMT.

We find that, consistent with expectations, lower-income households and those containing someone with driving limitations are more likely than others to own zero or fewer-than-expected vehicles and travel fewer miles. Among the segment of the population with higher income and no driving limitations, households that own zero or fewer-than-expected vehicles, and have lower VMT, tend to be more diverse, have fewer children, and live in rental units in very high density neighborhoods. The inclusion of attitudinal variables improves the ability to explain household vehicle ownership by a limited, but not trivial, amount. Individuals with more pro-environmental attitudes and who like transit, biking and walking are more likely to live in zero-vehicle-owning households. Conversely, those who like driving and living in spacious homes with large yards are more likely to be in higher-than-expected vehicle-owning households.

With respect to land use characteristics, both local density and regional status (the latter being a three-part classification based on metro-area size and the presence/absence of rail) yield strong associations. Households in higher-density neighborhoods are more likely to own zero or fewer-than-expected vehicles, and have lower VMT. But even lower-density living is associated with lower VMT if located in larger metropolitan areas (especially those with rail). Similarly, residential locations in smaller regions are found to have lower VMT if residential neighborhoods are denser. Overall, density has non-linear effects on travel behavior: a given increase in density is associated with larger reductions in households' VMT in lower-density neighborhoods than in higher-density ones, and this difference (between lower- and higher-density neighborhoods) is larger in smaller regions. Specifically among higher-density neighborhoods, however, the strongest relationships between density and VMT are found in large regions with rail: a given density increase is associated with larger reductions in VMT for households living in rail-served regions, all else equal.

The study provides useful insights for promoting the adoption of more sustainable travel behavior. In particular, it improves the understanding of what policy levers could lead to the adoption of environmentally-beneficial behaviors and help meet the required reductions in VMT. It also highlights the importance of collecting attitudinal data to improve the ability to explain vehicle ownership and use, especially decisions (such as voluntary "down-shifting" of vehicle ownership and VMT) that cannot be explained by traditional socioeconomic and demographic variables.

EXECUTIVE SUMMARY

Motivation of the Study

California has set ambitious targets for the reduction of greenhouse gas (GHG) emissions from transportation. There is wide consensus that improving vehicle technology per se will not allow meeting the established environmental targets, and other types of interventions that reduce car travel are needed. This includes promoting an array of transportation options that effectively fulfill the mobility needs of Californians, and designing policies to meet the required reduction targets for vehicle-miles traveled (VMT) that are as cost-effective as possible.

This research supports that goal through examining the factors influencing the transportation emissions footprint of Californians, and providing insights into the characteristics of small-footprint households. The aim of this project is to improve the understanding of the factors that lead some Californians to own fewer vehicles and travel less by car, thereby increasing the knowledge base to support policies that can promote such environmentally-beneficial choices.

Data

In this study (Part A), we investigate the impact of individual characteristics, e.g. household demographics and individual attitudes, as well as geographic location and urban form, on vehicle ownership and usage decisions. We analyze data from the 2009 National Household Travel Survey (NHTS). To maximize the sample size as well as the number and diversity of zero-vehicle-owning households available for this study, we use the NHTS dataset for the entire United States, and apply the Iterative Proportional Fitting method to weight the data to represent California's population on six key dimensions: household size, number of workers, number of household vehicles, household income, race/ethnicity, and population density.

Because a major limitation of the NHTS data is its lack of relevant attitudinal variables, we complement its analysis with one of attitudinal and behavioral data originally collected for four separate research projects in California between 1998 and 2011 by the co-PI (Mokhtarian)'s research team at UC Davis, and pooled for the purposes of the present study. For the NHTS data the unit of observation is the household, while for the attitudinal dataset it is the individual. Through the various stages of the research, we investigate the influence that household demographics, income and mobility constraints, individual attitudes and preferences, and residential neighborhood (land use) characteristics have on households' vehicle ownership and vehicle-miles traveled (VMT). As some of the attitudinal and behavioral datasets available from previous projects did not include information on VMT, the VMT-related portion of the analysis was conducted using only the NHTS data.

Outline of the Methodology

To study the variables associated with the choice of a small-footprint lifestyle, we first needed to establish what it means to "own fewer/more vehicles than expected". Accordingly, each household in the dataset(s) was classified as lower-than-expected (LTE), about-as-expected (AAE), or higher-than-expected (HTE) vehicle-owning, based on the comparison of the actual vehicle ownership level with the expected value computed from a model that was estimated for the project and that predicts vehicle ownership based on household size and composition (number of workers and drivers, and presence of children). The rationale is that these variables are the most fundamental markers of the "need" for a certain number of automobiles. We classified each household as AAE if the actual number of vehicles owned was within 0.5 (in either direction) of the expected number predicted by the model, and as LTE or HTE if the actual number of vehicles was, respectively, more than 0.5 lower or higher than the expected number. Households that did not own any vehicles were classified as zero-vehicle-owning (ZVO).

We next investigated the factors associated with these vehicle ownership *categories* (ZVO, LTE, AAE, and HTE). In particular, we were interested in exploring the reasons for which a household would own no vehicles, or fewer than expected vehicles. These reasons include income/ cost, driving limitations, residential location characteristics, and attitudes (e.g. dislike of driving, or various lifestyle orientations), among others. We incorporated variables representing these reasons into models of *vehicle ownership category*, in three stages:

- First, we wanted to account for the most natural *constraints* that could explain low or zero vehicle ownership status, namely **income and driving limitations**. These variables were available in both the NHTS and attitudinal datasets.
- Second, having taken the most common constraints into account, we wanted to focus on reasons for the *voluntary choice* of low/zero vehicle ownership status, namely **attitudes**. These variables were only available in the attitudinal datasets.
- Third, we reserved analysis of the role of **residential location** till last, in view of the methodological challenges associated with that variable. This analysis was conducted only on the NHTS dataset.

A similar approach was tested to model households' vehicle-miles traveled (VMT). However, given the relatively low explanatory power of the multinomial logit models that explained household VMT categories (i.e., whether households had LTE, AAE, or HTE vehicle-miles traveled), we instead decided to model VMT directly rather than model the VMT categories. Accordingly, we estimated a set of log-linear regression models of household VMT. Because several of the original attitudinal datasets from the previous projects did not contain information on household VMT, the estimation of VMT models was carried out only for the NHTS dataset:

- We first estimated a model that included only household composition - specifically the numbers of drivers and/or workers, and the presence of children (mirroring what was done for the definition of the vehicle ownership categories). We then incorporated income and driving limitations in the VMT model, as these are important predictors of car travel.
- To incorporate land use characteristics into the model, we created two variables: local density and regional status. For each household's residential location, a *local density* score was computed as a composite of four density-related characteristics, and the neighborhood was classified as *higher* or *lower* density, based on whether its score was greater or less than the mean. *Regional status* comprised three categories, consistent with the NHTS variable CBSACAT: *smaller regions*, i.e. metropolitan statistical areas (MSAs) with total population of less than one million, and non-MSAs; *larger regions without rail*, i.e. MSAs with population greater than one million not served by rail (subway); and *larger regions with rail*, i.e. MSAs with population more than one million that are served by rail. By interacting the local density and regional status variables, we created a land use typology of six neighborhood types.

- We next included land use characteristics in the model, testing different model specifications. We initially estimated a pooled model using the entire NHTS dataset and controlling for the residential density of the neighborhoods where the respondents live. Then, we estimated a segmented model that allows the model coefficients to vary by the six neighborhood types. Finally, we estimated a sample-selection model of VMT, which accounts for the impact of residential self-selection. It is important to note that while the impact of urban density was controlled for through the direct inclusion of that variable in most models (even those segmenting on neighborhood type), the influence of public transit on travel behavior was only accounted for in this study in an indirect way, through separating large MSAs with rail from those without rail and from smaller areas in the definition of regional status and thence in the estimation of the best models. Accordingly, it is not possible to quantify the relative contributions of density versus the presence of transit in influencing travel behavior toward greater sustainability. Indeed, even if the level of transit service were accurately quantified and included in the models, it would tend to be highly correlated with density and therefore it would still be difficult to distinguish their separate influences.

Key Findings

- Consistent with expectations, lower-income households, and (to a much lesser but still statistically significant degree) those containing someone with driving limitations, are more likely to own zero or fewer-than-expected vehicles. Similarly, they tend to have lower vehicle-miles traveled (VMT). All else equal, the more people in the household, the higher the VMT (*model results, NHTS and attitudinal dataset*).
- After accounting for household composition, income, and driving limitations, the inclusion of attitudinal variables improves the ability to explain household vehicle ownership by a limited, but not trivial, amount (specifically, provides a 12.2% increase in the explanatory power of the model) (*model results, attitudinal dataset*).
- Individuals with more pro-environmental attitudes and who like transit, biking and walking are more likely to live in zero-vehicle-owning households. At the other end of the spectrum, individuals who like driving and like to live in spacious homes with large yards are more likely to live in higher-than-expected (HTE) vehicle-owning households (*model results, attitudinal data*).
- Households belonging to the lower-than-expected (LTE) category or the zero-vehicle-owning (ZVO) category tend to be more ethnically/racially diverse, and more often live in rental units in high-density neighborhoods. They tend to have lower VMT and lower person-miles traveled (PMT) (*descriptive statistics, NHTS data*).
- These findings confirm that, in the general population, most households that do not own any vehicles do so *out of necessity*, because they have either limitations on their ability to drive or (far more often) low enough incomes to limit their ability to own vehicles (*descriptive statistics, NHTS data*).
- Among the households with higher income and no driving limitations (i.e. who have more space for voluntary changes in vehicle ownership), no large income differences are observed across vehicle ownership categories. This indicates that, beyond a certain income threshold, vehicle ownership decisions are largely made *out of choice*, and affected by non-income

variables such as residential location, individual attitudes and lifestyle preference (*descriptive statistics, NHTS data*).

- In particular, ZVO households with higher income and no driving limitations have comparable incomes to the households in the other vehicle ownership categories, but they are much more diverse, tend to live in smaller households with fewer children (i.e. have higher income per capita), more often live in rental units in very high density neighborhoods, and drive fewer miles thanks to the increased accessibility of central locations. The average population density of the neighborhoods where higher-income ZVO households live is more than four times the population density of HTE households' neighborhoods (*descriptive statistics, NHTS data*).
- Individuals who prefer (a) transit over driving, (b) biking and walking over driving, as well as (c) having shops within walking distance of their homes, are more likely than others to live in households with LTE vehicle ownership (*descriptive statistics, attitudinal data*).
- Higher-income individuals without driving limitations who live in ZVO or LTE households (1) are more likely to have attitudes supportive of a lower carbon footprint; (2) tend to have more such attitudes in combination; and (3) tend to hold those attitudes more strongly, compared to the rest of their peers. This indicates the value of policies directed at influencing pro-sustainability attitudes, and suggests that it may take the combined effect of several such attitudes to change behavior – only holding one such attitude but not others may not suffice (*descriptive statistics, attitudinal data*).
- Households that live in higher-density neighborhoods are more likely to be in the ZVO or LTE categories, and have lower VMT. The "worst-case" neighborhood type analyzed (lower-density neighborhood in a smaller region, i.e. with less than one million population) is associated with annual VMT that is substantially higher than the best-case (higher-density neighborhood in a larger region with rail) type (*model results, NHTS data*).
- Both local density and regional status matter: even lower-density living can be associated with lower VMT if located in larger metropolitan areas (especially those with rail), and even smaller regions can have lower VMT if residential neighborhoods are denser (*model results, NHTS data*).
- As income increases, households become more similar to the highest-income households in their propensity to own vehicles (or not). However, the convergence between wealthy and less-wealthy households happens from different directions depending on the interaction between regional status and residential neighborhood density:
  - In lower-density neighborhoods, as regional status *diminishes* (from larger region with rail to larger region without rail to smaller region) the less-wealthy become similarly likely to the wealthy to own cars (mostly *out of necessity*);
  - In higher-density neighborhoods, as regional status *increases* the wealthy become similarly likely to the less-wealthy to *not* own cars (mostly *out of choice*) (*model results, NHTS data*).
- The relationship between density and VMT is stronger (i.e. larger reductions in VMT associated with an increase in density) in lower-density neighborhoods than in higher-density ones. This difference is more pronounced in smaller regions (*model results, NHTS data*).
- However, among higher-density neighborhoods, the strongest relationship between density and VMT is found in large regions with rail: density increases are associated with larger

reductions in VMT for households living in these areas compared to other areas, after controlling for the impacts of other variables (*model results, NHTS data*).

- The study supports the importance of accounting for individuals' attitudes when studying vehicle ownership and travel behavior as they constitute important motivations for individuals' voluntary choices regarding vehicle ownership and mobility patterns (*descriptive statistics and model results, attitudinal data*).
- To increase the voluntary choice to reduce vehicle ownership, there is value both in modifying land use patterns as well as in trying to influence individuals' attitudes, preferably in combination (*descriptive statistics and model results, attitudinal data*).
- In future travel surveys, it would be desirable to collect appropriate information on individuals' attitudes to substantially improve the ability to explain (and interpret) the complex behaviors associated with vehicle ownership and use, especially those (such as voluntary "downshifting" of vehicle ownership and VMT) that cannot be explained by traditional socioeconomic and demographic variables alone.

1. Introduction

California has long been a leader in transportation policies to reduce smog-forming and particulate matter emissions, and more recently has positioned itself as a leader in policies to reduce greenhouse gas emissions from transportation. These policies have made important progress in reducing ambient air pollution and improving human health, and have the potential to help address the imminent threat of global climate change. Yet much remains to be accomplished. California has set a goal of an 80 percent reduction in greenhouse gases by 2050, with a significantly longer outlook than the federal government uses in federal policy development. To reach such an ambitious target requires a suite of regulations to reduce multiple pollutants, induce innovation in vehicle technology, and promote environmentally-beneficial behavioral changes, while at the same time targeting those regulations to be as cost-effective as possible.

In particular, the Sustainable Communities and Climate Protection Act (SB 375) in 2008 introduced the requirement for Metropolitan Planning Organizations (MPOs) to create *Sustainable Community Strategies* (SCSs) in order to meet established targets for GHG emissions in 2020 and 2035. As required by SB375, the California Air Resources Board (CARB) has set MPO-specific targets for GHG reductions. The state's MPOs are required to develop comprehensive plans for land use and transportation development, the Sustainable Community Strategies, to be integrated in their Regional Transportation Plans (RTP). The objective of a region's SCS is to meet the transportation and housing needs of its population while ensuring an appropriate reduction in the environmental impact from transportation and an increase in the livability of California's communities.

Achieving the targeted environmental goals is not simple. The recent California Advanced Clean Cars rulemaking exemplifies the California Air Resources Board's approach: use the very best research available to support the development of innovative and effective policies to improve ambient air quality and reduce greenhouse gas emissions. This study was designed to provide results from cutting-edge research that can support the staff at the Air Resources Board in the development of the next set of approaches to addressing the challenging issues of transportation emissions.

There is wide consensus that improving technology per se will not allow meeting California's ambitious environmental targets. Other types of interventions that reduce car travel will be needed, while at the same time promoting an array of transportation options that effectively fulfill the mobility needs of Californians. In this research, we examine the key factors influencing households to adopt, or inhibiting them from adopting, low-emissions travel patterns. We also examine other characteristics of small-footprint households (those with lower-than-average vehicle ownership and use). The study suggests leverage points that may be used to lower the barriers to low-emissions travel. The results of this work can help find ways to speed the transition to a world of low-emissions transportation to meet our mobility needs.

1.1 Research Objective

The aim of this project is to improve the understanding of the factors that lead some Californians to a small transportation emissions footprint (i.e. to own fewer vehicles and travel less by car), thereby increasing the knowledge base to support policies that can promote such environmentally-beneficial choices.[1] In this study, we investigate the impact of individual characteristics (e.g. household characteristics and individual attitudes and preferences) as well as geographic location and urban form on low transportation emissions households. We analyze data from the 2009 National Household Travel Survey and from several other California-specific surveys undertaken by research teams at UC Davis in order to build a comprehensive understanding of the contributions of the various factors to the formation of a low-emissions transportation footprint, while mitigating any limitations from the use of each data source.

In the study, we classify households as zero-vehicle-owning (ZVO), lower-than-expected (LTE), about-as-expected (AAE), or higher-than-expected (HTE) vehicle-owning, and perform a statistical analysis to identify the factors (1) leading to households falling into the different classifications, and (2) affecting the households' vehicle-miles traveled (VMT). The impacts of individual attitudes and preferences as well as residential location and geography on these choices are of particular interest in this analysis. Accordingly, to the extent possible, analyses are conducted with and without the inclusion of attitudinal measures, to enable the assessment of the improvements in the ability to explain observed choices associated with the inclusion of these variables. Similarly, we investigate the impact of residential location and geography, including methods to account for the self-selection of individuals into neighborhoods that match their residential preferences.

1.2 Background

A number of studies have analyzed how vehicle ownership and VMT vary with several individual and household characteristics. Anowar et al. (2014) provide a systematic overview and assessment of the methodological alternatives that have been introduced to model vehicle ownership, distinguishing between exogenous models that predict vehicle ownership as a function of other individual, household and location variables, and endogenous models that investigate the complex dynamics existing between vehicle fleet size and composition and vehicle usage, among other variables. Vehicle ownership is modeled in the academic literature and planning models through a variety of approaches, which can variously account, depending on the modeling purposes, for the demand and supply side of the car market, long- or short-run vehicle sales, and/or car-type segmentation, and are based on different assumptions about the impacts of household income, car costs, driver's license holding, sociodemographic variables, attitudinal variables, and treatment of scrappage (De Jong et al., 2004).

For example, Bhat and Guo (2007) used data from the San Francisco Bay Area travel survey, US Census data, US 2000 Tiger files, and Public Use Microdata Sample (PUMS) data to produce models based on the San Francisco population, finding that households with a high number of active and senior adults, employed individuals, higher income, and who live in owned dwellings tend to have higher car ownership propensity. On the other end of the spectrum, single-parent and single-individual households, households with physically challenged individuals, and

[1] This Part A report focuses on the "Factors Related to Voluntary Choice of Low Vehicle Ownership and Usage". A separate Part B report focuses on "Empirical Estimation of Household Vehicle Purchase and Usage Decisions".

households residing in multifamily housing units or prevalently composed of non-Caucasian and non-African races tend to have lower car ownership propensity. Similarly, Cirillo and Liu (2013) used multinomial logit models to analyze vehicle ownership using the subsample of the 2001 and 2009 National Household Travel Survey (NHTS) data related to the Maryland population. Similar to Bhat and Guo's results, the authors concluded that households with higher income as well as households with a larger number of workers and drivers tend to own more vehicles (Cirillo and Liu, 2013). In addition to vehicle ownership models, they also produced VMT models, using regression analysis, confirming that household income and the number of workers and drivers in the household significantly contribute to household vehicle use. Furthermore, an increase in operating costs decreases VMT, and a higher education level for the head of the household is associated with higher VMT.

Relatively few studies have investigated the influence of attitudes on vehicle ownership. Choo and Mokhtarian (2004) examined the impact of attitude and lifestyle variables on the choice of vehicle type, for example finding that small car owners tended to be more pro-environmental and pro-high-density than others, while large car and minivan owners tended to be less pro-high-density. Heffner et al. (2007) explored the adoption of newer technologies in vehicle choice in California, focusing in particular on consumers' intention towards the adoption of hybrid electric vehicles, and the role of social meanings and personal meanings (i.e. attitudinal factors that are not usually controlled for in traditional vehicle ownership or vehicle type choice models) in affecting the purchase of such vehicles. Cao et al. (2007) found that car dependence and a residential preference for outdoor spaciousness increased the propensity to own more cars, while a preference for accessibility decreased the propensity.

Delbosc and Currie (2012) investigated the reasons that cause some households to have low vehicle ownership, distinguishing between "involuntary" one-car households and "voluntary" one-car households. The study highlighted the extent to which the latter often live in more accessible locations, and do not experience restrictions on their mobility even if their average car travel is significantly lower than that of other households. However, involuntary low-vehicle owning households usually do not have access to as many transportation options, and usually rely more heavily on car-based travel, thus facing greater restrictions on their activities, with negative results on their social interactions and psychological well-being. The paper cautioned about the development of policies that attempt to limit car ownership if households cannot adjust to the negative consequences of the lower vehicle ownership through alternative transportation options.

Finally, Potoglou and Susilo (2008) discuss the use of ordered models, i.e. those modeling the number of vehicles owned by a household as a progressive number (1, 2, 3, etc.), vs. unordered models, i.e. those which consider each number of owned vehicles as an independent category and model the choice of a level of vehicle ownership based on the evaluation of the utility associated with that outcome. The authors compare the use of multinomial logit, ordered logit, and ordered probit car ownership models, using NHTS data and other datasets. They conclude that the multinomial logit model brings larger advantages in modeling the level of household vehicle ownership. Accordingly, this is the approach selected in our project to model vehicle ownership.

A large number of studies have been conducted and various methodologies employed to estimate the impact of individual, household and geographic variables on VMT. Williams et al. (2016) provide a comprehensive discussion of the methodologies used to estimate and forecast VMT for travel demand forecasting and planning purposes, and discuss the recent trends in VMT and the factors affecting changes in VMT in the United States. Blumenberg and Pierce (2012) explored the travel behavior patterns of lower-income households through the analysis of the 2009 NHTS data. The study confirmed that lower-income households are less likely to own cars and more likely to travel by non-car alternative modes. The analysis of the determinants of travel among lower-income categories highlighted how low-income adults rapidly convert rising income into additional mobility, at faster rates than for higher-income adults. Increased vehicle ownership is always associated with higher personal-miles traveled (PMT). However, the magnitude of this relationship is found to be larger among low-income adults.

Again, relatively few studies explore the influence of attitudes on VMT; two exceptions are Mokhtarian et al. (2001), and Handy et al. (2005). Mokhtarian et al. (2001) found that being an adventure seeker increased the use of a personal vehicle for short-distance trips (less than 100 miles one way), while a pro-environmental attitude reduced it. Handy et al. (2005) found that weekly vehicle-miles driven was positively influenced by a car dependent attitude and a residential preference for outdoor spaciousness, and negatively influenced by pro-bike/walk and pro-transit attitudes.

Finally, through a comprehensive review of the literature, Salon et al. (2012) systematically analyze the impact of 14 different factors on VMT. They quantify how much VMT can be expected to change in response to changes in specific policies or land use patterns, and discuss the likely impact (or lack thereof) of a number of VMT determinants on individuals' travel choices. The study highlights the large variation in the magnitude of the findings available from the literature and the need for better data sources to accurately quantify the impacts of changes in VMT.

The initial tasks in this project examine the impacts of socioeconomic traits and attitudes on vehicle ownership and VMT. But it is also important to incorporate land use variables into models of vehicle ownership category and VMT, for at least two reasons. First, geography matters to both of those outcomes: the evidence is strong that households located in denser urban areas, well-served by transit, tend to own fewer vehicles and travel less by car, all else equal (Cao et al., 2009). Second (and because of the first), understanding the influence of the built environment on VMT is essential to the successful implementation of policies that aim to affect travel behavior through modifications in the land use features, as in the case of SB 375. This landmark California legislation sets regional targets for greenhouse gas emission reductions due to VMT decreases, which in turn are expected to follow from changes in land use patterns arising from the legislatively-mandated Sustainable Communities Strategies developed by each region. Accordingly, it becomes more important than ever to have an accurate sense of the impact of land use characteristics on VMT. Given the emphasis of this study on household travel, the most appropriate land use characteristics to consider are those around the household's residential location.

Introducing residential location into the analysis adds considerable complexity, however, because housing choices are often endogenous with transportation choices. That is, a household may locate in a dense urban neighborhood, or near the workplace of the primary wage-earner, precisely so as to adjust its travel patterns, e.g. reduce its vehicle ownership or VMT for any of the reasons previously mentioned. Alternatively, it may happen to end up at such a location for other reasons, and only then decide it can get along with fewer cars and lower VMT. The built environment is a true causal influence only in the latter case; treating it as an explanatory variable in the former case may well overstate the influence of land use on transportation choices. For example, a household might be “enticed” to move to a dense urban neighborhood – not because it is consciously committed to reducing its vehicle ownership or VMT, but for reasons such as financial incentives (a possible policy instrument) or constraints on the supply of more-preferred housing. It is unlikely that such a household would reduce its auto travel as much as would the household who moves there specifically because it wants to do that very thing (Schwanen and Mokhtarian, 2005). Thus, to the extent that households self-select into residential locations consistent with their transportation preferences, those preferences may be the factors primarily responsible for the observed travel behavior, not the built environment per se. In reality, both the built environment and the attitudinal predispositions matter, and it is important to understand the extent to which each is operative in any specific context.

Addressing this “residential self-selection” issue is a key area of current research. A number of approaches to accounting for residential self-selection have been identified in the literature (Mokhtarian and Cao, 2008), each with advantages and disadvantages. Incorporating attitudes into an equation for vehicle ownership (or VMT) that already contains land use variables is one approach – referred to as “statistical control” – which is reasonably easy to understand and relatively effective. However, the National Household Travel Survey (although it contains a number of pertinent land use variables) does not contain data on many attitudes relevant to this particular issue. Conversely, the attitudinal dataset used in this project contains numerous appropriate attitude variables, but (1) for the most part, they include measures of vehicle ownership but not VMT, and (2) their land use measures are rather diverse and not always present for every case.

Ideally, we would like to have a dataset that is rich in both attitudinal and land use variables, so that we can compare the statistical control approach to other methods of controlling for residential self-selection that do not involve attitudes (since most datasets do not include them). Several methods of accounting for residential self-selection, including the use of (1) instrumental variables, (2) sample selection, and (3) propensity score and simultaneous equation approaches, do not require attitudinal variables to be explicitly measured (Bhat and Eluru, 2009; Boer et al., 2007; Zhou and Kockelman, 2008).

An approach to accounting for residential self-selection is taken by Salon (2015), who modeled VMT as a function of land use and socioeconomic variables, through the joint estimation of residential location type choice (selection) and location-type-specific VMT models in a sample-selection approach to control for residential self-selection. By incorporating a correction term into the VMT model for each location category (the six clusters, in our study), this approach controls for the endogeneity bias caused by the omission of attitudes that influence both residential choice and travel behavior. In this study, we use a similar approach to control for self-selection. Because we are not aware of a study that has employed a multinomial selection model with a multinomial outcome model, we apply this approach only to VMT, and not to vehicle ownership.

To do this, we estimate a joint sample-selection model that includes a multinomial logit model of residential location, and six VMT log-linear models (each one for each neighborhood type) which include the sample selection bias correction term that accounts for the likelihood of a given household to live in such a neighborhood type. Several types of selection bias corrections have been proposed in the literature. In the estimation of the sample-selection model for this project, we use the Lee correction term (Bourguignon et al., 2007).

1.3 Technical Approach

Using existing datasets, we develop a market segmentation analysis of zero- and low-vehicle/VMT households, through the development of the five project tasks (Task A.1 to Task A.5) described in the following sections.

We used two main types of data in this project, in complementary ways:

1. The 2009 National Household Travel Survey (NHTS) (http://nhts.ornl.gov/) is a large-sample dataset containing 150,147 households, including a national sample of 25,000 completed households and separate samples from twenty add-on areas that together added 125,147 completed households, with an overall response rate of 19.8% for the entire sample. The 2009 NHTS continues the series of household travel surveys begun by the Department of Transportation in 1969. In earlier waves, the travel survey was administered as the Nationwide Personal Transportation Survey, which became the NHTS with the 2001 data collection. The 2009 NHTS is the most recent dataset available from this program (a new wave is being administered in 2016 at the time this final report was written). The dataset is originally weighted to be representative of the U.S. population (FHWA, 2011). The California subset of the data is also large (N = 21,225 households before filtering out incomplete cases), and could have been the object of the analysis for this project. However, to have a larger sample available (particularly to increase the representation of zero-vehicle-owning households), we used the entire U.S. dataset, weighting it to be representative of the California population on six key variables (and further controlling for some of those and other variables by including them in the vehicle ownership and VMT models).[2] The NHTS has a wealth of information that is important to this project, including nine variables on mobility limitations as well as all the expected socioeconomic characteristics. In addition to vehicle ownership, it has several relevant measures of VMT: a detailed trip record for a single day, a self-reported estimate of annual VMT, and a computed "best estimate of annual miles" variable. However, the major drawback to the NHTS data is its lack of attitudinal variables relevant to our analysis. This drawback was addressed through the use of another dataset, as described below.
2. As part of previous projects conducted at UC Davis, the co-Principal Investigator of this study (Prof. Mokhtarian) and her colleagues have administered a number of surveys which are generally rich in attitudinal variables. They involve smaller samples (N < 3000 in each dataset) which are all California-based, albeit not representative of the state population as a whole. The four attitudinal surveys that were considered most useful to the current project are the Mobility Attitudes (Northern California, 1998), Caltrans Residential Location (Northern California, 2003), Fix I-5 Wave 3 (Sacramento Area, 2009), and Multitasking (Northern California, 2011) surveys. The pooled sample associated with these four surveys contains 8,024 cases, mainly concentrated in the urban areas of Northern California. Additional details about these four datasets are provided in Section 4. As described there, these attitudinal datasets provide a very useful complement to the large-

[2] The use of weights to make the dataset more representative of the population of interest would have been needed even if only the California subsample of the NHTS were used, as the NHTS data overrepresent the residents of the metropolitan areas in the state.

sample representative surveys such as NHTS, with respect to the vehicle ownership portion of the study. However, they will not allow an analysis of VMT, as this information was missing from most of these datasets.

Table 1.1 summarizes the number of vehicles, vehicles per person, vehicles per driver, vehicles per household and vehicle miles traveled in the United States in the years in which the NHTS and the four attitudinal datasets were collected. As confirmed by the numbers in the table, vehicle ownership and VMT did not change very much during these years (although some reduction in per-capita vehicle ownership has been observed after the mid-2000s, as discussed in Circella et al., 2016).

Table 1.1: Number of Vehicles, Vehicles per Person, Vehicles per Driver, Vehicles per Household and Vehicles Miles Traveled in the United States in the Years the Attitudinal Datasets Were Collected

Year	Vehicles (thousands)	Vehicles per person	Vehicles per driver	Vehicles per household	Vehicle miles driven (millions)
1998	203,169	0.737	1.098	1.982	2,417,852
2003	222,857	0.768	1.136	2.003	2,655,987
2009	234,468	0.764	1.119	2.001	2,633,248
2011	233,761	0.750	1.104	1.950	2,650,458

Source: Sivak (2015), using data from FHWA and U.S. Census Bureau.

The project (Part A) was divided into five main tasks (with the sixth task being the preparation of this final report):

- ***TASK A.1***: This task focused on the preliminary preparations required to use the NHTS data, and on classifying the households as *zero-*, *lower-than-expected*, *about-as-expected*, or *higher-than-expected* vehicle-owning (where the designations are relative to the expected number of vehicles that the "typical" household of that composition would own). We also performed a similar analysis for VMT, although in that case we ultimately rejected the classification approach in favor of continuing to predict VMT directly (rather than the VMT categories).
- ***TASK A.2***: To help identify the households for whom low vehicle-owning and low-VMT status is a lifestyle choice rather than an involuntary condition, in this task we evaluated the extent to which an NHTS household's vehicle ownership category, and VMT, could be explained by the constraints of household income and mobility limitations that restrict driving.
- ***TASK A.3***: To investigate the role of attitudes in leading individuals to make voluntary choices to reduce vehicle ownership, we turned to the attitudinal datasets. After replicating the relevant NHTS models on the attitudinal sample, we investigated the extent to which the inclusion of individual attitudes improved the predictive ability of the vehicle ownership category model already containing income and driving limitation variables.
- ***TASK A.4***: We further investigated the attitude-based reasons for some households to fall into the zero-vehicle-owning (ZVO) and lower-than-expected (LTE) vehicle ownership

categories, and analyzed other characteristics of the households identified as voluntarily choosing this lifestyle. This task provides further insight into the specific motivations of this segment, insight which is important to our ability to encourage greater adoption of the zero-/low-vehicle owning and low-VMT lifestyle.

- ***TASK A.5***: This task explored the role of geographic factors in a household's vehicle ownership and VMT status. Land use characteristics strongly influence vehicle ownership and VMT, and are the principal policy lever of SB375. This task is critical to ascertaining the extent to which land use itself is influencing vehicle ownership and VMT, as opposed to attitudinal predispositions prompting a household to locate within certain land use patterns. Failing to account for this residential self-selection process could lead to substantial overestimation of the role of land use in reducing VMT, with likely negative consequences for compliance with greenhouse gas emission reduction targets.

2. Task A.1: Classification of Households as Zero-, Low-, Medium-, or High-Vehicle-Owning

2.1 Preparation of the Dataset

In this part of the project, we used the 2009 NHTS data to investigate the reasons affecting a household's vehicle ownership and VMT. To do so, we performed a number of operations to clean and prepare the dataset, merge the original NHTS data files, aggregate various measures of VMT available in the NHTS data, and weight the dataset to be representative of the population of California.

To begin, we merged and processed information available from four different types of data files in the NHTS sample, which contained information respectively referring to the household, person, vehicle and trip levels. Although for this part of the project our unit of analysis is the household, this preliminary task was necessary because some important pieces of information (e.g. detailed information on each trip) were only available in data files at one of the other three levels. Further, because our unit of analysis is the household, we filtered out households with incomplete information for any member.

We then developed a process to properly weight the national NHTS sample to represent the California population. The goals of the process were to represent the California population in as much detail as is practical (given population data availability), while avoiding high variances of weights and maximum-to-minimum weight ratios (because extreme weights would produce undesirable effects on further statistical analysis performed on the dataset). We used the Iterative Proportional Fitting method (Kalton and Flores-Cervantes, 2003) to weight the data to be representative of California's population in 2011 (the last full year before the study began) with respect to several variables detailed below. We used expectation maximization[3] to impute missing data on household income, based on the other weighting variables.

We considered using the nationwide NHTS weights as a starting point for the development of new weights to represent California's population. Including the national weights would at least partially control for the effects of non-response bias (albeit for the national population) and for the over-sampling of specific portions of the population (in particular in the 20 regions for which additional cases were collected in the add-on areas of the 2009 NHTS - for more information, see FHWA, 2011) as well as race and ethnicity (although the distribution by race and ethnicity of the population of California is very different from that of the rest of the country). However, the exploration of this option revealed that starting from the national weights would have led to an undesirably high variance of the weights and maximum/minimum ratio of the weights. To avoid these negative results, we decided not to use the national NHTS weights, but to explicitly include

[3] In brief, as applied in this project, the expectation maximization method works as follows. In the first step, all missing data are imputed, perhaps by filling with sample means, but the imputed data are still tagged as such, to be distinguished from the non-missing data. The now complete dataset is used to calibrate models for predicting each variable as a function of all (or a subset of) the others. Then, those models are used to re-predict the imputed data points. With those updated values, the prediction models are re-calibrated and used to re-predict the imputed data points, and so on until the imputed data points do not materially change from one iteration to the next. For a more technical treatment of this method, see, e.g., de Waal et al. (2011).

race/ethnicity and population density as additional control targets in the Iterative Proportional Fitting process, which helped partially capture some of the desirable attributes of the national weights. In particular, because California's racial/ethnic profile is fairly unique compared to much of the United States, and because race and ethnicity are often important predictors of a number of socioeconomic variables, attitudes, and transportation-related choices, we considered it important to explicitly include them as weighting variables for California.

Similarly, we used information related to the "household location census tract average population density" to weight the data for population density and geography. For population density by census tract, we used the information from the 2010 decennial census to aggregate the number of households falling into each density category. Further, we condensed the eight categories from the census[4] (which are also available for the NHTS data) into three main categories to reduce the total number of controls used in the definition of the weights (again to help minimize large variations in the weights). We chose the following categories of population density, which approximately match the categories "urban", "suburban", and "rural" as originally defined by the Department of Defense when establishing this criterion for zip codes, and defined the census tract where a household lives as:

- rural, if the population density is below 1000 persons per square mile;
- suburban, if the population density is between 1000 and 3000 persons per square mile;
- urban, if the population density is above 3000 persons per square mile.

Thus, in the process that was used to determine the final set of weights for this dataset, we used Iterative Proportional Fitting to weight the raw (unweighted) NHTS data to replicate California's population distributions on these six key dimensions:

- household size;
- number of workers;
- number of household vehicles;
- household income;
- race and ethnicity; and
- population density.

By this we mean that the finally-weighted NHTS sample matches the distribution of California's 2011 population on each of these variables taken individually. In general, it will not match the population on every possible combination of those variables. For example, the share of weighted cases in the $50,000 - $74,999 income category will be the same as the share in California's population at large, and the share of cases in the suburban density category will match California, but the share of suburban households with $50,000 - $74,999 income will not necessarily match. In some cases, information on the joint distribution of two or more variables is not readily available. In other cases, it is possible to match the joint distribution of two variables at a time, but (again to prevent large variations in the weights) we limited the use of two-way crosstabulations, and only controlled for the following joint distributions in the Iterative Proportional Fitting application:

[4] The initial eight categories that were considered are: 0-99; 100-499; 500-999; 1,000-1,999; 2,000-3,999; 4,000-9,999; 10,000-24,999; and more than 25,000 persons/square mile.

- household size by number of workers;
- number of workers by number of household vehicles; and
- number of household vehicles by household size.

Thus, the weighted shares of cases in the sample *will* match those of California statewide with respect to these pairs of variables in combination.

Finally, to prepare the variables needed for this project, we aggregated information on the various VMT variables for each household using information from the various NHTS files:

- self-reported annual VMT (from the NHTS *person* file);
- best estimate of annual VMT (from the NHTS *vehicle* file);
- travel day diary VMT and PMT by transportation mode (from the NHTS *trip* file).

The information about VMT from detailed trip data was processed in order to account for the higher occupancy of vehicles for trips made by carpooling with other travelers, and to avoid double-counting the household VMT for trips made by more than one household member who traveled together in the same vehicle for some trips. All analyses in this report were conducted on the weighted sample, unless explicitly indicated otherwise.

2.2 Creating Vehicle Ownership Categories

After completing the preparatory task of appropriately weighting the NHTS data to represent the distribution of households in California on the dimensions of interest the first major task of the project was to establish, for the purposes of this study, what it meant to "own fewer (or more) vehicles than expected". Our premise was that the need to own a certain number of vehicles is most fundamentally based on household composition, specifically on variables associated with the number, ages, employment, and driver's license status of household members. Accordingly, as explained below, in this task, after separating out the zero-vehicle-owning (ZVO) households, we classified the remaining households in the 2009 NHTS data as lower-than-expected (LTE), about-as-expected (AAE), or higher-than-expected (HTE) vehicle-owning, based on the comparison between the actual level of vehicle ownership in the household and the vehicle ownership level that would be expected for that household given its current size and characteristics.

We considered it important to keep the ZVO category separate, and not to combine it with households owning some number of vehicles, in the LTE category. The identification of zero-vehicle-owning (ZVO) households is straightforward: there are 6,562 of them in the unweighted NHTS sample that was used in this project – about 5% of the unweighted sample, which corresponds to 10,458 households in the weighted sample, comprising 8.0% of the total households in California. For the remaining cases, the strategy was to (1) develop a multinomial logit model[5]

[5] Multinomial logit is the most common model used for cases in which the dependent variable takes on discrete values, reflecting the decision-maker's choice among discrete alternatives. Each alternative is characterized by a utility function (generally a linear function of observed variables, plus an error term representing the combined impact of all unobserved variables), the individual is assumed to choose the alternative with the highest utility (which is unknown to the analyst because of the unobserved influences on utility), and the model provides the estimated probability that each alternative will be chosen, given the observed characteristics of that alternative and of the individual. For a technical introduction to multinomial logit and other discrete choice models, see Ben-Akiva and Lerman (1985).

of vehicle ownership for which the choice alternatives were 1, 2, 3, or 4+ vehicles, and for which the only explanatory variables were those associated with household characteristics; (2) use the model to predict vehicle ownership for each case in the estimation sample; and (3) classify each case as LTE, AAE, or HTE depending on whether its vehicle ownership was lower than, about the same as, or higher than the model's prediction.

Table 2.1 shows the results of the multinomial logit model for vehicle ownership categories, where 4 or more vehicles is used as the base category. The model has a reasonably good ρ^2 goodness-of-fit value of 0.3440, using the equally-likely (EL) model as base (i.e. the set of explanatory variables included in the model improves the log-likelihood of the model by a large amount, compared to the "null" model that assigns an equal likelihood of being chosen to all alternatives)[6]. Approximately 48.9% of the cases were correctly classified by this model. The specification of the model balanced the need to increase the ability to explain household vehicle ownership against the need for parsimony in the number of variables included in the model (this balance is captured by the adjusted ρ^2 goodness-of-fit value, which penalizes for the number of parameters estimated in the model).

The estimated coefficients show that the more driving workers there are in the household, the less likely the household is to own fewer vehicles. In other words, and as expected, households with more driving workers tend to own more vehicles. On the other hand, the presence of non-driving adults (either working on or non-working) is associated with higher likelihood of a household to own fewer vehicles. Similarly, the presence of children in a household is associated with higher likelihood of owning more vehicles. For the latter two variables, we initially tested the *number* of people in each category, as with the first two variables. However, better results were obtained by using dummy variables for the *presence* of any people in each category. The implication is that once the first two categories (driving workers and driving non-workers) are accounted for, and given the presence of one person in either of the latter two categories, the addition of more people in that category has only a negligible influence on the number of vehicles owned by a household.

---

[6] Similarly (but not identically) to the R^2 measure in regression, ρ^2 for a discrete choice model can be interpreted as the proportion of information that is explained by the model, and ranges from 0 for the naïve equally-likely model, to 1 for the perfect model. ρ^2 values around 0.3 or better are considered relatively good for a disaggregate travel behavior model (Hensher et al., 2005). The equally-likely log-likelihood (LL_EL) is a goodness-of-fit benchmark for the most naïve model in which all alternatives are considered equally likely. The constants-only log-likelihood is a goodness-of-fit benchmark for the next-most naïve model (containing only constants), in which for everyone, the probability of choosing a given alternative is equal to the sample-wide market share of that alternative. The final-model log-likelihood (LL_f) is the comparable benchmark for the final model. Log-likelihoods will be negative, and the closer to 0, the better (the hypothetical perfect model, in which everyone's choices have predicted probabilities of 1, has a log-likelihood of 0). The ρ^2 measure represents the share of the log-likelihood distance between a naïve model and the perfect model that is covered by the final model, i.e. $\rho^2 = 1 - (LL\_f/LL\_EL)$.

Table 2.1: Estimated Coefficients and Goodness of Fit for the Best Vehicle Count Multinomial Logit Model (NHTS Dataset, Weighted N = 120,024)

	Vehicles			
	1	2	3	4+
Driving workers (#)	**-4.849** (-187.70)	**-2.204** (-113.09)	**-0.937** (-55.70)	(base)
Driving non-workers (#)	**-4.186** (-158.83)	**-1.881** (-89.89)	**-0.699** (-37.04)	(base)
Non-driving workers (#)	**1.007** (12.82)	**0.412** (5.73)	**0.156** (2.13)	(base)
Presence of non-driving non-workers (DV)	**0.313** (12.90)		**0.164** (5.86)	(base)
Presence of children (DV)	**-0.174** (-8.78)		**0.182** (10.18)	(base)
Constant	**10.455** (191.54)	**6.564** (135.83)	**2.828** (61.38)	(base)
Log-likelihood final model	-109147.78			
Log-likelihood constants only	-147945.62			
Log-likelihood equally-likely model	-166388.59			
ρ^2 (equally-likely base)	0.3440			
Adjusted ρ^2 (EL base)	0.3439			

*Note: coefficients are in bold and t-statistics are reported in parentheses. All estimated coefficients are significant at least at the 0.1% level, unless otherwise noted. Coefficients marked with * are significant at the 5% level.*

To assign households to vehicle ownership categories (lower than expected, about as expected, higher than expected), we used the (household-specific) predicted probabilities from the multinomial logit model to calculate the expected number of vehicles owned by the household, from the following formula:

$$E = 1\times Pr[1] + 2\times Pr[2] + 3\times Pr[3] + 4\times Pr[4+] ,$$

where E is the weighted average of the possible numbers of vehicles a household could own (simplifying “4 or more” to “4”), and the weights are the respective probabilities of owning each number of vehicles.

The aim is then to classify a household as

- Lower-Than-Expected (LTE) vehicle-owning, if the actual vehicle ownership is on the low side of what the model predicts;
- About-As-Expected (AAE) vehicle-owning if the actual vehicle ownership is close to that predicted by the model; and
- Higher-Than-Expected (HTE) vehicle-owning if the actual ownership is on the high side of the model prediction.

The question is how to define "on the low side", "close", and "on the high side". Because there is a certain amount of noise in the equation for household vehicle demand (even households of identical composition, residential location, income, and attitudes may appropriately have a different demand for vehicles because of differing activity patterns), we experimented with different "caliper widths" for this classification, including 0.5, 0.75, and 1.00. Ultimately, we concluded that 0.50 was the most logical cutoff, representing the standard round-up versus -down threshold. Accordingly, we used a threshold of 0.5 below/above the expected number of owned vehicles, to define the vehicle ownership category of a household. Specifically, defining the residual (R_n) number of vehicles as "the actual number of vehicles minus the expected number of vehicles for household *n*", the household was classified as:

- Lower-Than-Expected (LTE) if $R_n \leq -0.5$;
- About-As-Expected (AAE) if $-0.5 < R_n < 0.5$; and
- Higher-Than-Expected (HTE) if $R_n \geq 0.5$.

In addition to the three categories above, a household that does not own any vehicles was classified as "zero-vehicle-owning", or ZVO.

Figures 2.1 and 2.2 illustrate the distribution of households in the various vehicle ownership categories used in this project in the unweighted and weighted datasets, respectively. It can be seen that weighting increases the effective presence of the otherwise underrepresented ZVO and LTE households in the sample, and conversely diminishes the size of the HTE group.

Figure 2.1: Distribution of Households by Vehicle Ownership Category (Unweighted NHTS Dataset, N=130,474)

<table>
<tr><th colspan="2" rowspan="2"></th><th colspan="6">Expected</th><th rowspan="2">Total*</th></tr>
<tr><th colspan="6">1 1.5 2 2.5 3 3.5 4</th></tr>
<tr><td rowspan="4">Actual</td><td>1</td><td>30,739</td><td colspan="5">6,982</td><td>37,721</td></tr>
<tr><td>2</td><td>7,198</td><td colspan="2">43,508</td><td colspan="3">1937</td><td>52,643</td></tr>
<tr><td>3</td><td colspan="3">16,276</td><td colspan="2">5,675</td><td>218</td><td>22,169</td></tr>
<tr><td>4+</td><td colspan="5">10,258</td><td>1,121</td><td>11,379</td></tr>
<tr><td colspan="8">ZVO: 6,562 LTE: 9,137 AAE: 81,043 HTE: 33,732</td><td>123,912</td></tr>
</table>

*Note: the total in this column does not include the 6,562 households that are in the *zero-vehicle-owning* (ZVO) category

Figure 2.2: Distribution of Households by Vehicle Ownership Category (Weighted NHTS Dataset, N=130,482)

<table>
<tr><th colspan="2" rowspan="2"></th><th colspan="4">Expected</th><th rowspan="2">Total*</th></tr>
<tr><th>1 – 1.5</th><th>1.5 – 2 – 2.5</th><th>2.5 – 3 – 3.5</th><th>3.5 – 4</th></tr>
<tr><td rowspan="4">Actual</td><td>1</td><td>32,065</td><td colspan="3">10,640</td><td>42,705</td></tr>
<tr><td>2</td><td>5,486</td><td>39,449</td><td colspan="2">3,621</td><td>48,556</td></tr>
<tr><td>3</td><td colspan="2">10,641</td><td>7,350</td><td>438</td><td>18,429</td></tr>
<tr><td>4+</td><td colspan="3">7,809</td><td>2,525</td><td>10,334</td></tr>
<tr><td colspan="2"></td><td>ZVO: 10,458</td><td>LTE: 14,699</td><td>AAE: 81,389</td><td>HTE: 23,936</td><td>120,024</td></tr>
</table>

*Note: the total in this column does not include the 10,458 households that are in the *zero-vehicle-owning* (ZVO) category

2.3 Defining VMT Categories

During the development of the project, we tested the creation of VMT categories based on household structure-related variables for VMT using a process analogous to the one used to create the categories for vehicle ownership. However, since the dependent variable, VMT, is continuous rather than discrete, we used linear regression models instead of multinomial logit models to predict VMT. The goal was to create essentially the same categories for VMT as for vehicle ownership: Zero- (ZVM), Lower-Than-Expected (LTE), About-As-Expected (AAE), and Higher-Than-Expected (HTE) VMT. However, it developed that since the NHTS did not collect VMT information for zero-vehicle households, there were relatively few households with zero VMT. Accordingly, we decided to classify all zero-VMT households as LTE.

The NHTS dataset contains several measures of household VMT. One can be computed from the trip files, representing the VMT created by a household on the single travel diary day of the survey. Another can be obtained by summing over vehicles the self-reported annual miles driven by each vehicle in the household. A third variable, BESTMILE, combines both these sources of information (together with other information about the household) to create what is considered the best estimate of the household's annual VMT, and accordingly that is the measure we used. We log-transformed the VMT measure BESTMILE to bring the distribution of the dependent variable closer to normal, in keeping with the standard regression assumptions, and we added 1 to all values to prevent taking the log of 0 (thereby mapping zero-VMT households to a value of 0 on the transformed scale). To deal with heteroscedasticity, we used weighted least squares (WLS) to estimate a linear regression model for ln(VMT+1). With respect to explanatory variables, we used the same subset of variables as for the vehicle ownership category determination models:

- number/presence of driving workers;
- number/presence of driving non-workers;

- number/presence of non-driving workers;
- number/presence of non-driving workers;
- number/presence of children; and
- lifecycle category of household.

We estimated a linear regression model of ln(VMT+1), using the explanatory variables listed above, and used the estimated coefficients from this model to predict the expected VMT for each household in the NHTS dataset (through saving the values from the initial regression model, and transforming them back from ln(VMT+1) to VMT). However, with the continuous dependent variable VMT, defining LTE, AAE, and HTE categories was not as clear-cut as it was for the discrete vehicle ownership variable: how close would actual VMT need to be to the predicted value, to consider it to be "about as expected"? We considered five different tolerance levels for the difference between predicted and actual VMT: 10%, 15%, 20%, 25%, and 30% of the actual household VMT. Ultimately, we chose a 20% threshold, meaning that actual VMT was AAE if predicted VMT differed from it by no more than 20% in either direction, and LTE or HTE if predicted VMT were more than 20% higher or lower, respectively.

Due to the low goodness of fit of the models estimated using the VMT categories as the dependent variable, as discussed in Section 3.2, in the remainder of the project we did not use the VMT categories further. Instead, in the analysis of the factors that affect a household's VMT, we focused on the direct estimation of log-linear regression models of the household's VMT.

3. Task A.2: Predicting Vehicle Ownership Category and Vehicle-miles traveled as a Function of Household Income and Driving Limitations

3.1 Vehicle Ownership

Having predicted, in Task A.1, what a household's vehicle ownership category "should" be, based on its size and characteristics, and having classified it as zero-vehicle-owning (ZVO), lower-than-expected (LTE), about-as-expected (AAE), or higher-than-expected (HTE), it is useful to investigate the factors determining how a household is classified.

In particular, we are interested in why a household would own no vehicles, or fewer than expected. Asking the questions "why drive?" and "why not drive?" suggests the following most-common explanations:

- *Why drive?*
    - I enjoy the act of driving
    - I have little choice, given geography + desired activity pattern
    - I don't enjoy exercise
    - Mobility limitation on using transit or walking/biking
    - Driving is just easier/more convenient
    - Takes less time than alternative
    - *Etc.*

- *Why not drive?*
    - I don't enjoy driving
    - Pro-environmental lifestyle choice
    - Pro-urban lifestyle choice
    - Pro-exercise lifestyle choice
    - Physical/mental limitation on ability to drive
    - License revoked
    - Another alternative is just easier/more convenient
    - I can't afford it, can't obtain insurance
    - *Etc.*

The main reasons for being a ZVO or LTE household, then, can be categorized as income/cost-related, driving limitations-related, attitude-related (e.g. dislike of driving, or the various lifestyle orientations), or what we might call happenstance convenience. The first two types of these reasons – both relating to constraints – are available in the NHTS dataset. The other two types – relating to voluntary choice – are not.[7]

[7] Note that several kinds of attitudes could lead to ZVO or LTE status, not just a pro-environmental one. Somebody might simply want to get exercise by walking/biking whenever possible, for example, or might like living "close to the destination they need to reach" or "where the action is". These attitudes may well be correlated with a concern for the environment, but are not identical to it. We explore the impact of personal attitudes in the following Task A.3.

Accordingly, we used the NHTS data to investigate how much explanatory power the first two types of reasons have in determining ZVO or LTE status. Specifically, taking the dependent variable to be the vehicle ownership category assigned in Task A.1 (rather than the number of vehicles per se), which takes on the values LTE, AAE, and HTE, together with the "zero-vehicle-owning" (ZVO) category, in this task we develop a multinomial logit model containing only income and mobility limitations as explanatory variables.

We tested many different possible multinomial logit model specifications, particularly exploring the explanatory power of various indicators of mobility limitations, such as the need to use special transit services or the inability to use bus/subway. The model with only a driving limitation dummy variable (indicating the presence in the household of anyone with a medical condition that precludes driving), together with income, yielded the best results. Household income was measured with five categories ($0 to 24,999, $25,000 to 49,999, $50,000 to 74,999, $75,000 to 99,999, and $100,000 or more), and controlled for in the models through a set of four dummy variables, using $100,000 or more (highest income category) as the base category. Table 3.1 summarizes the estimation results for the final multinomial logit model of vehicle ownership category that was estimated in this task. This model has a ρ^2 (equally-likely (EL) base) goodness-of-fit measure of 0.2857, indicating that it explains almost 28.6% of the information in the data (which is considered relatively good for a disaggregate model of travel behavior, particularly a discrete choice model with four alternatives and so few explanatory variables; Hensher et al., 2005).

Approximately 46.2% of the cases were correctly classified by this model. The results from this best model are quite logical: lower-income households, and those containing someone with driving limitations, are more likely than others to own zero or fewer-than-expected vehicles. Between these two explanatory variables, income was by far the more important predictor. The multinomial logit model with only income as the explanatory variable has a ρ^2 (equally-likely (EL) base) goodness-of-fit measure of 0.2787, which means that adding the driving limitation variable improved the ρ^2 of the vehicle ownership category model only by 0.0070 (or 2.5% of the total goodness of fit). This is probably due to the fact that only a small fraction, roughly 7.5% of the households, have a member with a driving limitation in the NHTS dataset, and the vehicle ownership of these households is not very different from that of other households. Another explanation is that a household member with a driving limitation is likely not to have a driver’s license, and the impact of non-driving adults on vehicle ownership was largely accounted for in the baseline model estimating the actual number of vehicles owned by the household (Task A.1, Section 2). In other words, the number of vehicles a household would be expected to own had already been adjusted downward if the member with a driving limitation did not have a license (and even more so if not working), so that such a household may be classified as owning about as many vehicles as expected, rather than fewer than expected – leaving little for the driving limitation variable to additionally explain at this stage. Nevertheless, the driving limitation variable *is* statistically significant in this model, and does add a small amount of explanatory power to the model.

Table 3.1: Estimated Coefficients for the Best Vehicle Ownership Category Multinomial Logit Model (NHTS Dataset, Weighted N = 130,329)

	ZVO	LTE	AAE	HTE
Income Level 1 (Less than $25,000)	**2.591** (57.99)	**0.436** (16.62)	(base)	**-0.901** (-36.38)
Income Level 2 ($25,000 to 49,999)	**1.325** (27.52)	**0.312** (12.07)	(base)	**-0.521** (-24.82)
Income Level 3 ($50,000 to 74,999)	**0.353** (5.89)	**0.116** (4.00)	(base)	**-0.205** (-9.77)
Income Level 4 ($75,000 to 99,999)	**0.192*** (2.73)	**0.087*** (2.65)	(base)	**-0.109** (-4.72)
Income Level 5 ($100,000 or larger)	(base)	(base)	(base)	(base)
Driving Limitation (DV)	**1.373** (50.16)		(base)	
Constant	**-3.721** (-88.51)	**-1.913** (-102.49)	(base)	**-0.930** (-73.87)
Log-likelihood final model	-129033.07			
Log-likelihood constants only	-137063.64			
Log-likelihood equally-likely model	-180674.36			
ρ^2 (EL base)	0.2857			
Adjusted ρ^2 (EL base)	0.2856			
*Note: coefficients are in bold and t-statistics are reported in parentheses. All estimated coefficients are significant at least at the 0.1% level, unless otherwise noted. Coefficients marked with * are significant at the 1% level.*				

3.2 Vehicle-miles traveled (VMT)

Using a process similar to the one used to estimate the vehicle ownership category multinomial logit models, we estimated multinomial logit models of the VMT categories defined in Task A.1, including only the income dummy and driving limitation variables. Income was by far the most important predictor; the driving limitation variable was never statistically significant in any of these VMT models. However, since the best multinomial logit model for household VMT category performed rather poorly in terms of goodness of fit, with a ρ^2 measure never exceeding 0.0359, we decided to focus on modeling the continuous ln(VMT+1) variable directly, from this point forward.

Therefore, we estimated weighted least squares (WLS) log-linear regression models for the continuous dependent variable ln(VMT+1), beginning with a model which included only information on household composition, and specifically the numbers of drivers and/or workers, and the presence of children in a household (mirroring what was done for vehicle ownership, in which we first controlled for the influence of household composition, though in that case in an indirect way through the influence of those variables on classifying households into vehicle ownership category). This model had an R^2 of 0.2872. Table 3.2 summarizes the coefficients that were estimated for this model. The p-values for all variables in the model are very small, indicating that all estimated coefficients are strongly significant, which is not surprising for a large sample comprising 119,413 (weighted)[8] cases. Given the large sample sizes for all model estimations using the NHTS data, the level of statistical significance is usually very strong, and most estimated coefficients in the final NHTS models are statistically significant at least at the 1% level, unless otherwise specified.

Table 3.2: Estimation Results for the ln(VMT+1) Model (NHTS Dataset, Weighted N = 119,413)

Explanatory Variable		Beta	S.E. Beta	P-Value
Number of Driving Workers		0.6672	0.0032	<0.0001
Number of Driving Non-Workers		0.4219	0.0039	<0.0001
Number of Non-Driving Workers		-0.0640	0.0118	<0.0001
Number of Non-Driving Non-Workers		-0.0328	0.0060	<0.0001
Number of Children under the Age of 16		0.1451	0.0030	<0.0001
Constant		8.5115	0.0064	<0.0001
R^2	0.2872			
Adjusted R^2	0.2872			

The results from the model estimation are quite logical: the presence of driving workers in the household tends to increase VMT more than the presence of driving non-workers, which is consistent with expectations due to the increased needs for commuting trips of driving workers and due to the higher incomes associated with having workers in the household. The numbers of non-drivers (either workers or non-workers) in the household are associated with a small negative effect on household VMT. The negative signs are probably due to an unobserved variable bias:

[8] As mentioned earlier, since the NHTS did not collect VMT information for zero-vehicle-owning (ZVO) households, the households belonging to this group were not included in the VMT analysis, reducing the sample size to 119,413 weighted cases.

both these variables are negatively correlated with household income, and their negative sign in this model seems to reflect the influence of lower income on a household's VMT level, rather than capturing a "true" effect of these variables on VMT. Finally, and not surprisingly, households with more children under the age of 16 also tend to have higher VMT, although the impact on household VMT of the presence of an additional child below 16 is lower than the impact of the presence of an additional driving adult in the household.

Next, we added four dummy variables for the income categories (using the highest income category as the reference group in the model estimation) and a dummy variable indicating the presence of one or more household members with driving limitations; the results are reported in Table 3.3 below. This model has an R^2 of 0.3204 (a 12% increase from the first model), and all estimated coefficients are statistically different from zero, apart from the coefficient for the number of non-driving workers. Almost the entire increase in the goodness of fit of this model is associated with the impact of the income variables rather than of the driving limitation dummy variable. Excluding the latter from the group of explanatory variables would cause a drop in the R^2 of only 0.003, but we retain it because it is statistically significant, conceptually relevant, and explains some, even if only a small amount of, variation in VMT.

Table 3.3: Estimation Results for the Best ln(VMT+1) Model (NHTS Dataset, Weighted N = 119,408)

Explanatory Variable	Beta	S.E. Beta	P-Value
Income Level 1 (Less than $25,000)	-0.5541	0.0078	<0.0001
Income Level 2 ($25,000 to 49,999)	-0.3328	0.0071	<0.0001
Income Level 3 ($50,000 to 74,999)	-0.1221	0.0072	<0.0001
Income Level 4 ($75,000 to 99,999)	-0.0578	0.0080	<0.0001
Driving Limitation (DV)	-0.0943	0.0116	<0.0001
Number of Driving Workers	0.5632	0.0034	<0.0001
Number of Driving Non-Workers	0.3923	0.0038	<0.0001
Number of Non-Driving Workers	0.0034	0.0116	0.7665
Number of Non-Driving Non-Workers	0.0515	0.0065	<0.0001
Number of Children under the Age of 16	0.1332	0.0029	<0.0001
Constant	8.8570	0.0084	<0.0001
R^2	0.3204		
Adjusted R^2	0.3203		

The results from this best WLS ln(VMT+1) model are quite logical: lower-income households, and those containing someone with driving limitations, tend to have lower VMT, all else equal. After controlling for the impact of income, all household composition variables have the expected positive coefficients (i.e., the more people in the household, the higher the VMT). Further, regardless of working status, the effect of the number of drivers in the household on VMT is larger than the effect of the number of non-drivers. Among non-driving individuals, non-workers are associated with higher household VMT, probably because of the increased escorting needs for non-commuting trips.

4. Task A.3: Including Attitudinal Variables in the Vehicle Ownership Category Model

4.1 Preparation of the Dataset

Having accounted for the major *constraints* that influence household vehicle ownership and VMT, we next wanted to investigate the attitudes that may motivate a household to *voluntarily* reduce its vehicle ownership. Since the NHTS survey does not have many usable attitudes, a different dataset, consisting of the pooled samples obtained from four different surveys administered across different years (1998 to 2011) and locations in California in connection with previous research projects developed at the University of California, Davis, was used for this purpose. As these surveys did not collect VMT data, this task could only be conducted on vehicle ownership. As background to the remainder of the discussion of this dataset, it is important to note that the unit of analysis is the individual, whereas for the NHTS sample it is the household.

A number of initial activities needed to be performed to assemble the dataset that was used in this part of the analysis. An initial list of seven different datasets previously collected in California was considered. The process of pooling the data available from the different sources involved tradeoffs between sample size and data consistency. On one hand, we wanted the dataset to contain the largest possible number of cases. On the other hand, we wanted the variables in the dataset to be (1) present and (2) consistently measured, for as many of the constituent samples as possible. Only in some cases did a variable in one dataset have an exact counterpart in all other datasets. In other cases, variables were "similar" across datasets, but they were not measured in exactly the same way. This most often occurred with attitudinal variables. For example, an attitudinal factor that measured a person's pro-environmental disposition was sometimes based on different individual "pro-environmental" items. In such cases, it was a judgment call whether to consolidate the different measurements into the "same" attitudinal variable, or to consider them as separate, or process them further, or possibly omit one or more datasets completely, in order to increase consistency across the remaining datasets.

After comparing the information available for each sample, the datasets collected from the following four projects were used in this part of the research:

1. Mobility Attitudes (Northern California, 1998)
2. Caltrans Residential Location (Northern California, 2003)
3. Fix I-5 Wave 3 (Sacramento Area, 2009)
4. Multitasking (Northern California, 2011)

Table 4.1 summarizes key information on these four studies. The first two surveys were mailed to randomly-selected residential addresses in the judgmentally-selected study neighborhoods (chosen to represent urban and suburban locations). Specifically, the Mobility Attitudes survey was mailed to 8,000 respondents of three neighborhoods in the San Francisco Bay Area. With an overall response rate of more than 25%, after discarding responses with too much missing data we retained about 1,900 cases that were useful for the purposes of this study (Mokhtarian et al., 2011). The invitations to complete the Caltrans Residential Location were sent to 8,000 addresses in eight different neighborhoods of Northern California. As there were only 6,746 valid addresses in the sample and the number of responses totaled 1682, the response rate for this survey was 24.5% (Handy et al., 2005). The Fix I-5 survey was conducted entirely online. Because of the specific characteristics of this survey, and the multiple channels through which the invitation to complete the survey was distributed, it is not possible to compute a proper response rate for this study (Ye et al., 2012). Finally, the Multitasking survey was available in both paper and online forms. It also was distributed through multiple channels in 2011 among the residents of 16 counties in Northern California (Neufeld and Mokhtarian, 2012). The pooled sample associated with the four surveys considered most useful to the current project contains 8,024 cases, mainly concentrated in the urban areas of Northern California.

Table 4.1: Summary Information for the Four Attitudinal Datasets Used in the Study

Year data collected	1998	2003	2009	2011
Location	Northern California: Pleasant Hill, Concord, and North San Francisco	Northern California (8 neighborhoods located in Santa Rosa, Sacramento, Modesto and the Silicon Valley)	Sacramento area	Northern California (16 counties in Bay Area and Sacramento)
Sample size (for present study)	1,904	1,217	2,054	2,849
Sampling approach	Random selection of residential addresses	Random selection of residential addresses	State agency employees, commute alternatives listserv, Fix I-5 information listserv, large employers	Transit riders, carpool permit holders, transportation management associations, UC Davis staff, online panels, random residential addresses, etc.
Response rate	>25%	24.5%	N/A	N/A
Survey type	Paper	Paper	Online	Paper and online
Back-ground reference	Mokhtarian et al. (2001)	Handy et al. (2005)	Ye et al. (2012)	Neufeld and Mokhtarian (2012)

The final sample obtained from merging the four attitudinal datasets above contained 8,024 cases. Although some of these datasets are relatively old, their interest for the present study lies in the information they can provide on relationships among individual attitudes, household characteristics and vehicle ownership. We believe that these relationships are relatively robust over time. Since we were not interested, in this study, in measuring the impact of other variables (e.g. the adoption of technology) which might be more subject to changes over time, the utility of increasing the sample size of the dataset used for this part of the analysis prevailed over the importance of having more recent information.

Extensive data processing was performed to create new variables from those available, and to maximize the comparability of variables across the four individual datasets. After importing all the relevant variables into this new dataset, the expectation maximization algorithm was used to impute missing data for variables such as number of vehicles, household income, household size, total number of workers in the household, and some attitudinal variables. The final combined dataset includes attitudinal variables related to several dimensions such as pro-environment, pro-transit, pro-high (urban) density, pro-biking/walking, and pro-driving, among others, in addition to all socioeconomic variables such as income, household size, and so on.

Similar to what was done in Task A.1 for the NHTS dataset, we needed to create weights to make the attitudinal datasets representative of the California population. Although the datasets were collected from previous projects in California, they are not representative of the entire population in the state. Accordingly, in the interest of making the analyses of NHTS and the attitudinal datasets as comparable as possible, we applied a similar procedure to the one developed in Task A.1 to replicate the distribution of California's population on household size, number of vehicles, number of workers, and income in the attitudinal dataset. We did not control for race and ethnicity and for population density in weighting the attitudinal dataset, because these variables were not available for all samples that were used in this task.

4.2 Vehicle Ownership

Similar to the track for the NHTS dataset (Section 2.2), after classifying as ZVO any household not owning any vehicles, we classified the remaining households as LTE, AAE, or HTE vehicle-owning by estimating an multinomial logit model for the number of vehicles in the household, and then comparing the actual number of vehicles owned by the household with the expected number of vehicles predicted by that model. In this way, we created the vehicle ownership category variable (with values ZVO, LTE, AAE and HTE) that was used as the dependent variable in subsequent multinomial logit models. In the attitudinal dataset we were unable to subdivide the drivers and non-drivers in terms of their worker/non-worker status as we could for the NHTS sample. Instead, we included a drivers × workers interaction term to account as well as possible for non-linear effects of these two variables. Table 4.2 shows the results of the multinomial logit model for number of vehicles in the household estimated with the attitudinal dataset, where 4 or more vehicles is used as the base category. This model also has a reasonably good ρ^2 value of 0.3482 (using the equally-likely (EL) model as base), which is comparable to that of the analogous model estimated with the NHTS data (Table 2.1).

Next, just as for the NHTS database, we estimated a multinomial logit model of household vehicle ownership category that included only income and driving limitations. This model (results not shown) has a ρ^2 of 0.2499 for the combined attitudinal dataset, which is similar to the goodness of fit of 0.2857 of the comparable model estimated with the NHTS data in Task A.2 (Table 3.1). The model estimated with the attitudinal dataset is able to correctly classify 44.1% of the cases (compared to 46.2% for the NHTS data). The interpretation of the model is as expected, with the presence of a driving limitation as well as lower income contributing to the individual's household owning fewer vehicles.

Table 4.2: Estimated Coefficients for the Best Vehicle Count Multinomial Logit Model (Attitudinal Dataset, Unweighted N = 7,715)

	Vehicles			
	1	2	3	4+
	-1.588	**-0.402**	**-0.216**	
Workers (#)	(-15.11)	(-7.64)	(-4.25)	(base)
	-3.851	**-1.169**	**-0.344**	
Drivers (#)	(-38.22)	(-18.81)	(-6.17)	(base)
	0.500			
Drivers × workers (#)	(12.33)			(base)
	9.771	**5.164**	**2.054**	
Constant	(47.18)	(33.04)	(13.83)	(base)
Log-likelihood final model	-6971.42			
Log-likelihood constants only	-9081.78			
Log-likelihood equally-likely model	-10695.26			
ρ^2 (equally-likely base)	0.3482			
Adjusted ρ^2 (EL base)	0.3472			

*Note: coefficients are in bold and t-statistics are reported in parentheses. All estimated coefficients are significant at least at the 0.1% level, unless otherwise noted. Coefficients marked with * are significant at the 5% level.*

We then began to add attitudinal variables to measure the improvement in the ρ^2 associated with each combination of variables. We tested several different model specifications that included attitudinal variables. Ultimately, the best multinomial logit model included eight attitudinal variables. The estimation results of this model, which has a ρ^2 of 0.2804 (compared to the EL base), are reported in Table 4.3. The sample size of the dataset used for the estimation of this model is about 4% smaller than 8,024 due to the presence of missing values for some of the attitudinal variables.

The interpretation of the model is again natural. Those who like the idea of using transit, biking and walking (each in its own right, without reference to other modes), and having shops within walking distance of home, are more likely to own zero or fewer vehicles than expected, compared to those with the opposite views. Similarly, those who like transit, biking, and walking *over driving* are also more likely to own zero or fewer vehicles than expected compared to those who do not. Those who like driving and/or large yards are more likely to own more vehicles than expected. Among the variables that were tried but that were not significant in the final model were the factor scores for the “pro-driving”, “pro-high density”, and “pro-transit” attitudes. Among the factor scores that were available in the attitudinal dataset, only the factor score for the pro-environment attitude was found to have a statistically significant effect and hence included into the final model. As expected, a higher factor score on the pro-environment dimension increases the probability of a household owning zero vehicles.

The ρ^2 value of 0.2804 represents an improvement of 0.0305, or a 12.2% increase, in the explanatory power of the model without inclusion of individuals’ attitudes. The inclusion of the attitudinal variables also leads to an increase of 1.2 percentage points (2.7%) in the share of cases correctly classified (44.1% for the model without attitudes, and 45.3% for the model with attitudes). These results indicate that attitudes are useful in explaining vehicle ownership choices, but with an incremental influence that is relatively small compared to that of income. It is likely that even greater predictive ability could be achieved by measuring a more customized set of attitudes, and measuring them consistently across cases, neither of which could be done with the sample at hand. Accordingly, it seems important for future large-scale travel-behavior surveys to include attitudinal questions in addition to the widely used socioeconomic questions.

A summary of the improvements in goodness of fit achieved with the addition of each block of variables to the multinomial logit models of household vehicle ownership category using the attitudinal dataset is reported in Table 4.4. As shown in the table, the models estimated with this dataset have satisfactory goodness of fits, which are comparable to those of the analogous models estimated with the NHTS data. In addition, the inclusion of individual attitudes increases the goodness of fit of the model by a limited but not trivial amount, and contributes to the incremental ability of the models to correctly predict choices.

Table 4.3: Estimation Results for the Best Vehicle Ownership Category Multinomial Logit Model (Attitudinal Dataset, Weighted N = 7,984)

	ZVO	LTE	AAE	HTE
Income Level 1 (Less than $25,000)	**4.090** (13.61)	**0.597** (7.84)	(base)	**-0.956** (-9.63)
Income Level 2 ($25,000 to 49,999)	**2.684** (8.73)		(base)	**-0.799** (-9.42)
Income Level 3 ($50,000 to 74,999)	**1.693** (5.10)		(base)	**-0.541** (-6.34)
Income Level 4 ($75,000 to 99,999)	**0.855*** (2.12)		(base)	**-0.311** (-3.41)
Income Level 5 ($100,000 or more)	(base)	(base)	(base)	(base)
Driving Limitation (DV)	**0.892** (7.19)		(base)	**-0.529** (-3.61)
Factor Score for Pro-Environment	**0.156*** (2.96)		(base)	
Like Driving			(base)	**0.079*** (2.43)
Like Shops within Walking Distance of Home	**0.500** (8.32)		(base)	
Large Yard	**-0.124*** (-2.97)		(base)	**0.271** (9.47)
Like Transit	**0.299** (4.86)		(base)	
Like Transit over Driving	**0.323** (6.04)	**0.176** (6.35)	(base)	
Like Biking and Walking	**0.267** (5.05)	**0.090*** (2.78)	(base)	
Like Biking/Walking over Driving			(base)	**-0.140** (-5.07)
Constant	**-9.479** (-20.49)	**-2.370** (-17.40)	(base)	**-1.498** (-7.64)
Log-likelihood final model	-7964.37			
Log-likelihood constants only	-8890.68			
Log-likelihood equally-likely model	-11068.27			
ρ^2 (equally-likely base)	0.2804			
Adjusted ρ^2 (EL base)	0.2782			

*Note: coefficients are in bold and t-statistics are reported in parentheses. All estimated coefficients are significant at least at the 0.1% level, unless otherwise noted. Coefficients marked with * are significant at the 5% level.*

Table 4.4: Summary of the Vehicle Ownership Category Models (Attitudinal Dataset)

Model Specification	Sample Size	ρ^2 (EL Base)	ρ^2 (MS Base)	Percentage of Cases Correctly Classified	% Improvement in ρ^2 (EL base) over Base Model	% Improvement in Cases Correctly Classified over Base Model	Percentage-Point Improvement in Cases Correctly Classified
ASCs only (Base Model)	8024	0.1973	0	41.7%	-	-	-
ASCs + Income only	8024	0.2462	0.0610	44.1%	24.8%	5.76%	2.4
ASCs + Income + Driving Limitations	8024	0.2499	0.0656	44.1%	26.7%	5.76%	2.4
ASCs + Income + Driving Limitations + 8 Attitudes	7978	0.2804	0.1042	45.3%	42.1%	8.63%	3.6

5. Task A.4: Classifying ZVO and LTE Vehicle Ownership Households on the Basis of Likely Reasons for their Status

The models estimated in the previous tasks A.2 and A.3 allowed us to predict the probability that a household owns fewer/more than the expected number of vehicles. However, for planning and practical purposes, it is of interest to further disaggregate the reasons for which households fit into some specific vehicle ownership categories vs. others. In particular, we are interested in investigating the reasons for which specific households belong to the ZVO and LTE categories, and especially the voluntary decisions associated with such vehicle ownership levels. To do this, we analyze both the NHTS sample and the attitudinal dataset to investigate the relationships behind a household being in a ZVO or LTE category, and explore the likely reason(s) for their status, focusing in particular on the role of driving limitations, income, and (for the attitudinal dataset) individual attitudes in affecting these choices.

We hypothesize there being a precedence *hierarchy* among those reasons, where driving limitations precede household income which precedes attitudes:

driving limitation → *household income* → *attitudes*. (5.1)

That is, if one is mobility-limited, he/she may be unable to drive regardless of income or attitudes, and if one has very low income, he/she may be unable to drive even if physically able and wanting to do so.

The group of individuals who do not have driving limitations or income constraints are of particular interest for this study. These individuals might voluntarily decide to reduce their auto ownership and not drive based on other factors, including their personal preferences and attitudes. Understanding the reasons for which some of the individuals in this group fall into the ZVO or LTE categories provides useful information about the causal mechanisms affecting the choices of those that have more *margin* to adjust their travel behavior and auto ownership, and therefore informs the development of policies that can encourage similar environmentally-beneficial choices.

5.1 Characteristics of Households in Lower Vehicle Ownership Categories (NHTS Data)

Table 5.1 identifies the main categories of users based on household income and driving limitations (to simplify the analysis, in this table, as well as in other parts of this task, we classify household income as simply “higher” or “lower”). The cell marked with a * denotes the households who have higher incomes and who do not include any member with driving limitations, who have more “space” for voluntary choices and who therefore are of particular interest for this part of the research.

This group is something of a “black box” for the NHTS dataset, because this dataset does not include information about personal attitudes. In fact, when households in this cell have lower than expected auto ownership, we expect it to be due to the impact of variables other than driving limitations and income, including the potential impact of geographic location and/or individual attitudes. Accordingly, later in this section we will investigate this group of cases (individuals) using the attitudinal dataset from Task A.3.

Table 5.1: Classification of Households by Driving Limitations and Household Income

		Household Income	
		Lower	*Higher*
Driving Limitations	*No*		*
	Yes		

Table 5.2: Classification of Households by Vehicle Ownership Category in the NHTS Data

	Unweighted		California Weighted^	
Value	N	%	N	%
ZVO	6,562	*5.0*	10,458	*8.0*
LTE	9,137	*7.0*	14,699	*11.3*
AAE	81,043	*62.1*	81,389	*62.4*
HTE	33,732	*25.9*	23,936	*18.3*
Valid Total	130,474	*100%*	130,482	*100%*
Missing	30*		22*	
Total	130,504		130,504	

Table 5.2 summarizes the distribution of households in the NHTS dataset (respectively in the unweighted data, and in the data weighted to be representative of the California population) by the four vehicle ownership categories ZVO, LTE, AAE, and HTE. A relatively low share of households (19% of the weighted sample) falls into the combined categories of zero-vehicle and lower-than-expected vehicle owning. This is not surprising, considering that more than three-fifths (62%) of households are in the "about as expected" category. It is also interesting to note that the households who were *sampled* (i.e., based on the unweighted NHTS data) are less likely to fall into the zero and lower-than-expected vehicle owning categories (12%) than households in the overall population, possibly highlighting the difficulty of obtaining a sample that is fully representative on these dimensions of interest to planners. In particular, survey respondents often tend to be better educated and have higher income than average.

Similarly to the other tasks of this project, in the remainder of this section we will first focus on the California-weighted NHTS data, and then move to the analysis of the attitudinal dataset available from the previous research projects at UC Davis. Tables 5.3 and 5.4 summarize the cross-tabulation of vehicle category respectively with income categories and with the presence of individuals with driving limitations in the household, using the California-weighted NHTS data (please note that small differences in the total number of cases are due to rounding issues in the weighted dataset). As expected, most households in the higher income categories tend to fall into the higher vehicle-owning categories, while households containing people with driving limitations are far more likely than others to own zero or fewer than expected vehicles.

Table 5.3: Cross-tabulation of Household Vehicle Ownership Category by Income Category in the NHTS Data (Weighted Dataset, N= 130,481)

		Vehicle Category				
		ZVO	LTE	AAE	HTE	Total
Annual Household Income	Less than $25,000	6,848 *23.7%*	3,629 *12.5%*	15,908 *55.0%*	2,550 *8.8%*	28,935 *100.0%*
	$25,000 to 49,999	2,127 *7.4%*	3,756 *13.0%*	18,624 *64.5%*	4,363 *15.1%*	28,870 *100.0%*
	$50,000 to 74,999	573 *2.6%*	2,399 *10.9%*	14,476 *65.5%*	4,653 *21.1%*	22,101 *100.0%*
	$75,000 to 99,999	325 *2.1%*	1,621 *10.4%*	10,067 *64.6%*	3,562 *22.9%*	15,575 *100.0%*
	$100,000 or more	585 *1.7%*	3,294 *9.4%*	22,313 *63.8%*	8,808 *25.2%*	35,000 *100.0%*
	Total	10,458 *8.0%*	14,699 *11.3%*	81,388 *62.4%*	23,936 *18.3%*	130,481 *100.0%*

Note: As a comparison, the median household income in California was $58,931 in 2009 (source: US Census, available online at https://www.census.gov/prod/2010pubs/acsbr09-2.pdf).

Table 5.4: Cross-tabulation of Household Vehicle Ownership Category by Driving Limitation Status in the NHTS Data (Weighted Dataset, N= 130,304)

		Vehicle Category				
		ZVO	LTE	AAE	HTE	Total
Driving Limitations	No	7,455 *6.2%*	13,720 *11.4%*	76,824 *63.7%*	22,531 *18.7%*	120,530 *100.0%*
	Yes	2,834 *29.0%*	979 *10.0%*	4,556 *46.6%*	1,405 *14.4%*	9,774 *100.0%*
	Total	10,289 *7.9%*	14,699 *11.3%*	81,380 *62.5%*	23,936 *18.4%*	130,304 *100.0%*

Table 5.5 summarizes the relationship between driving limitations and household income: as expected, the two variables are not independent, but individuals with driving limitations are more

often found to live in households with lower income, probably because of the causality relationships between these variables: individuals with driving limitations often have a lower ability to work and generate income, or caring for such an individual may limit the ability of the caregiver's ability to engage in paid employment. Please note that also in this table we grouped the household income categories into only two main - and therefore coarser - categories, namely "lower" income (defined as an annual household income below or equal to $50,000, before taxes) and "higher" income (above $50,000). The definition of "lower" household income is clearly somewhat subjective and depends on the local characteristics of a region (e.g. an individual's purchasing power is lower, all else equal, in regions where the cost of living is higher) and the characteristics of the households. Incidentally, this classification of household income divides it at approximately the median for the U.S. ($50,221 in 2009 - somewhat lower than the median household income in California, which was $58,931 in 2009 [9]). More importantly, however, this threshold was considered useful for the purposes of this study, namely to identify households in which income might play a stronger role in limiting vehicle ownership. It is reasonable to expect that households with annual income above $50,000 are less subject to income limitations (and have more space for voluntary choices regarding vehicle ownership). The highlighted cell in Table 5.5 includes the group of households that have high income and no driving limitations: this group of households will be the object of the following analyses, in which we will investigate the relationships behind vehicle ownership, with the purpose of exploring the *voluntary* choices of households that choose to own zero or fewer than expected vehicles.

Table 5.6 compares the average household characteristics and travel behavior between the households that fall into the lower (ZVO or LTE) vehicle ownership categories and those that belong to the higher (AAE or HTE) vehicle ownership categories, for the entire NHTS dataset. While the average household size and numbers of workers, drivers and children are similar, several differences in the other household characteristics emerge between the two groups. Households in the lower vehicle ownership categories tend to be more ethnically/racially diverse (in particular, they tend to have about double the percentages of Hispanics and Blacks compared to the higher-ownership group, but slightly fewer Asians), have lower income, and more often include members with driving limitations. Further, they more often rent their housing unit and live in much denser neighborhoods than the households in the higher vehicle ownership categories (AAE and HTE). Further, the households in the lower vehicle ownership categories, on average, travel less using private vehicles (their average daily person-miles traveled, or PMT, is much lower than households in the higher vehicle ownership categories) and have significantly lower daily VMT (27.6 vs. 53.0).

[9] Source: https://www.census.gov/prod/2010pubs/acsbr09-2.pdf (last accessed March 29, 2017).

Table 5.5: Classification of Households by Income Level and Driving Limitation Status in the NHTS Data (Weighted Dataset, N= 130,326)

			Driving Limitations		Total
			No	Yes	
Household Income	Lower	Number of Households	50,620	7,017	57,637
		% within Income Category	*87.8%*	*12.2%*	*100.0%*
		% within Driving Limitation Category	*42.0%*	*71.8%*	*44.2%*
		% of Total	38.8%	5.4%	44.2%
	Higher	Number of Households	**69,932**	2,757	72,689
		% within Income Category	***96.2%***	*3.8%*	*100.0%*
		% within Driving Limitation Category	***58.0%***	*28.2%*	*55.8%*
		% of Total	**53.7%**	2.1%	55.8%
Total		Number of Households	120,552	9,774	130,326
		% within Income Category	*92.5%*	*7.5%*	*100.0%*
		% within Driving Limitation Category	*100.0%*	*100.0%*	*100.0%*
		% of Total	92.5%	7.5%	100.0%

Table 5.6: Average Household Characteristics and Travel Behavior Indicators of Households in the Lower (ZVO and LTE) vs. Higher (AAE and HTE) Vehicle Ownership Categories (Weighted NHTS Dataset, N=130,482)

	Lower Vehicle Ownership Categories (ZVO or LTE, N=25,157)	Higher Vehicle Ownership Categories (AAE or HTE, N=105,325)
Household size	2.9	2.9
# Drivers	1.7	1.8
# Workers	1.1	1.2
# Children	0.4	0.4
Household income	$49,753	$69,821
% Hispanic	10.5%	5.6%
% Asian	11.5%	12.6%
% Black	11.9%	5.3%
% Other	3.7%	1.9%
Limitations on driving (DV)	15.3%	5.7%
% Owning housing unit	54.9%	85.0%
Residential density (housing units/square mile)	6,027	2,547
Rental units in neighborhood (%)	46.3%	32.1%
Population density (population/square mile)	10,769	6,187
Employment density (employees/square mile)	2,412	1,555
Daily PMT*	49.0	78.5
Daily VMT*	27.6	53.0
# Household vehicles	1.3	2.1

*These numbers are based on the travel diary information from the NHTS dataset, while the dependent variable used in the estimation of the VMT models is the BESTMILE estimate of a household's annual VMT provided in the NHTS dataset.

Table 5.7 provides a further disaggregation of these average household characteristics, through the comparison of households in all four vehicle categories that have been used so far in the project. Households that belong to the ZVO category, in particular, are found to have a much lower average income than the households in all other vehicle ownership categories. They have smaller household sizes, and much more often consist of minorities (in particular, they are less likely to include White/Caucasians or Asians, and much more likely to include Hispanics, Blacks, and Other races), include individuals with disabilities, live in rental units, and live in very dense neighborhoods. Not surprisingly, the members that live in these households tend to travel much less by motorized vehicles (average PMT and VMT for the households in the ZVO category are respectively only 10.3 and 3.4) than those who live in households with vehicles. These findings confirm that, in the general population, most households that do not own any vehicles appear to do so out of necessity, because they either have limitations on driving or have low income conditions that limit their access to vehicle ownership. Interestingly, the average characteristics of those in the LTE category tend to be closer to those of the AAE group than to those of the ZVO group.

Table 5.7: Average Household Characteristics and Travel Behavior Indicators of Households in Each Vehicle Ownership Category (Weighted NHTS Dataset, N=130,482)

	ZVO (N=10,458)	LTE (N=14,699)	AAE (N=81,389)	HTE (N=23,936)
Household size	2.0	3.5	2.8	3.2
# Drivers	0.7	2.4	1.8	1.9
# Workers	0.5	1.4	1.2	1.3
# Children	0.2	0.5	0.4	0.5
Household income	$ 33,578	$ 61,262	$ 67,329	$ 78,295
% Hispanic	12.4%	9.1%	5.8%	5.1%
% Asian	7.0%	14.8%	13.3%	10.2%
% Black	19.9%	6.2%	5.4%	5.0%
% Other	4.9%	2.8%	1.9%	1.6%
Limitations on driving (DV)	27.5%	6.7%	5.6%	5.9%
% Owning housing unit	32.3%	71.0%	82.7%	92.8%
Residential density (housing units/square mile)	8,187	4,490	2,715	1,976
Rental units in neighborhood (%)	55.5%	39.7%	33.1%	28.8%
Population density (population/square mile)	13,242	9,010	6,507	5,100
Employment density (employees/square mile)	2,851	2,100	1,638	1,274
Daily PMT*	10.3	70.6	72.8	97.5
Daily VMT*	3.4	41.0	49.1	66.0
# Household vehicles	-	1.3	1.8	3.1

*These numbers are based on the travel diary information from the NHTS dataset, while the dependent variable used in the estimation of the VMT models is the BESTMILE estimate of a household's annual VMT provided in the NHTS dataset.

Having analyzed the characteristics of all households in the population, it is of greater interest, for the purposes of our project, to focus on the subset of households with higher incomes that do not include any members with driving limitations. Learning more about the households in this category that chose to own zero or fewer vehicles than expected can provide useful information for promoting similar environmentally-beneficial behaviors among other groups of individuals. Accordingly, Table 5.8 reports the average characteristics for the 69,932 higher-than-median-income households that do not include members with driving limitations (as identified in Table 5.5), respectively classified in lower vs. higher vehicle ownership categories. Among the households that do not have income or driving limitations, differences in income and household composition between the households in the lower vs. higher vehicle ownership categories are rather small. In other words, above a certain income threshold, income does not seem to play an important role in explaining vehicle ownership. This supports the assumption of the precedence hierarchy of reasons affecting vehicle ownership that was presented at the beginning of this section: the differences in vehicle ownership and driving behaviors among this target group of households are likely to be explained by other factors, including geographic location and individual attitudes and preferences.

Table 5.8: Average Household Characteristics and Travel Behavior Indicators of the Subset of Higher-than-Median-Income Households with No Driving Limitations in Lower vs. Higher Vehicle Ownership Categories (Weighted NHTS Dataset, N=70,011)

	Lower Vehicle Ownership Categories (ZVO or LTE, N=8,352)	Higher Vehicle Ownership Categories (AAE or HTE, N=61,659)
Household size	3.4	3.2
# Drivers	2.3	2.1
# Workers	1.7	1.5
# Children	0.5	0.5
Household income	$93,905	$96,071
% Hispanic	5.1%	3.7%
% Asian	18.7%	16.0%
% Black	5.4%	3.6%
% Other	2.0%	1.3%
% Owning housing unit	75.0%	91.7%
Residential density (housing units/square mile)	7,391	2,494
Rental units in neighborhood (%)	42.0%	29.0%
Population density (population/square mile)	11,862	6,031
Employment density (employees/square mile)	2,627	1,517
Daily PMT*	67.4	95.4
Daily VMT*	39.9	64.3
# Household vehicles	1.4	2.4

*These numbers are based on the travel diary information from the NHTS dataset, while the dependent variable used in the estimation of the VMT models is the BESTMILE estimate of a household's annual VMT provided in the NHTS dataset.

Table 5.9 reports the average household characteristics for the same higher-than-median-income households with no driving limitations for all four vehicle-owning categories. Households in the lower vehicle ownership categories (LTE, and in particular ZVO) have, on average, an income level that is comparable to the income of households in the higher vehicle ownership categories. Still, households in the lower vehicle ownership categories more often belong to an ethnic/racial minority, more often live in rental units, and are more often located in the denser parts of cities. In particular, the average population density of the neighborhoods where the higher-income households who own zero vehicles live is more than four times the population density of the HTE households' neighborhoods. Further, and as expected, individuals in these lower vehicle ownership categories travel much less by motorized private vehicles than individuals in the higher VO categories.

Table 5.9: Average Household Characteristics and Travel Behavior Indicators of Higher-Income Households with No Driving Limitations in Each Vehicle Ownership Category (Weighted NHTS Dataset, N=69,920)

	ZVO (N=1,330)	LTE (N=7,021)	AAE (N=45,288)	HTE (N=16,281)
Household size	2.5	3.6	3.2	3.4
# Drivers	1.3	2.5	2.1	2.1
# Workers	1.4	1.7	1.5	1.5
# Children	0.3	0.5	0.5	0.6
Household income	$ 91,911	$ 94,283	$ 95,496	$ 97,669
% Hispanic	6.3%	4.9%	3.7%	3.9%
% Asian	13.7%	19.6%	17.6%	11.7%
% Black	11.0%	4.4%	3.5%	4.0%
% Other	2.3%	1.9%	1.3%	1.5%
% Owning housing unit	48.8%	80.0%	90.3%	95.6%
Residential density (housing units/square mile)	17,354	5,504	2,676	1,989
Rental units in neighborhood (%)	59.2%	38.7%	29.6%	27.1%
Population density (population/square mile)	21,453	10,045	6,361	5,114
Employment density (employees/square mile)	4,078	2,352	1,602	1,281
Daily PMT*	16.4	76.4	90.6	108.8
Daily VMT*	10.5	45.1	60.8	73.8
# Household vehicles	-	1.4	2.0	3.3

*These numbers are based on the travel diary information from the NHTS dataset, while the dependent variable used in the estimation of the VMT models is the BESTMILE estimate of a household's annual VMT provided in the NHTS dataset.

The profile of these "lifestyle" households in the ZVO category seems to match some of the emerging trends in transportation that have been reported in many studies (e.g. Circella et al., 2016): diverse, smaller households (average household size is only 2.5 for the ZVO category, compared to 3.2-3.6 for the other vehicle ownership categories) with few children (only 0.3 children/household, on average), rather high income and no driving limitations, who live in rental units in the central parts of cities, and drive relatively few miles thanks to the increased accessibility and transportation options available to the more central residential location. The stage in life (and related lifecycle effects, e.g. for college students and younger adults) could also play a role, even if it was not directly investigated as part of the current study. Lifecycle effects are likely to explain part of the motivations behind the ZVO status of households that are transiting through different stages. For example, ZVO households may become vehicle-owning households in the future as their lifecycle changes (Circella et al., 2016), and conversely for HTE households.

5.2 The Role of Individual Attitudes in Determining Vehicle Ownership Categories

The analysis of NHTS data is useful for estimating the proportions of the California population falling into each category of interest, and for investigating the travel behavior of the group(s) of interest to this study. However, except by indirect inference it cannot address the question of *why* a given household with ample income and no driving limitations would own zero or fewer-than-expected vehicles. For insight into that question, we turn to the analysis of the attitudinal dataset used in the previous Task A.3. For this analysis, it is again important to keep in mind that although the unit of observation in the NHTS sample is the household, for the attitudinal sample it is the individual.

Table 5.10 compares the distribution of cases (households in the NHTS dataset, individuals in the attitudinal dataset) by vehicle ownership categories in the NHTS and attitudinal datasets for both the unweighted data and the data that were weighted to be representative of the population of California. Note in particular that the representation of ZVO cases in the attitudinal dataset needed to be doubled during the weighting process, whereas LTE cases were already in approximately the proper proportion. The agreement between the distributions in the two datasets is relatively close, which suggests that insights can be reasonably transferable from one dataset to the other, despite the use of households in one case and individuals in the other.

Table 5.10: Classification of Households by Vehicle Ownership Category in the NHTS and Attitudinal Datasets

	NHTS Unweighted		NHTS California Weighted		ATT Unweighted		ATT California Weighted^	
Value	N	*%*	N	*%*	N	*%*	N	*%*
ZVO	6,562	*5.0*	10,458	*8.0*	309	*3.9*	643	*8.0*
LTE	9,137	*7.0*	14,699	*11.3*	1,186	*14.8*	1,158	*14.4*
AAE	81,043	*62.1*	81,389	*62.4*	4,834	*60.2*	4,683	*58.4*
HTE	33,732	*25.9*	23,936	*18.3*	1,695	*21.1*	1,540	*19.2*
Valid Cases	130,474	*100%*	130,482	*100%*	8,024	*100%*	8,024	*100%*
Missing	30		22		0		0	
Total	130,504		130,504		8,024		8,024	

^ Unlike the NHTS sample, all attitudinal datasets used for this pooled sample are California datasets – however, they still need weighting to properly represent the California population on roughly the same variables used to weight the NHTS data.

Tables 5.11 and 5.12 summarize the cross-tabulation of vehicle ownership categories respectively by income categories and the driving limitation status of the survey respondent, using the attitudinal dataset (N=8,024). Finally, Table 5.13 summarizes the distribution of respondents by coarser income categories and driving limitation status in this dataset. It can be seen that those with driving limitations comprise about 8% of the population. The households to which those individuals belong are far more likely than others to have annual incomes below $50,000 (74% versus 42%, respectively).

Table 5.11: Cross-tabulation of Household Vehicle Ownership Category by Income Category in the Attitudinal Dataset (Weighted Dataset, N= 8,024)

		Vehicle Category				
		ZVO	LTE	AAE	HTE	Total
Annual Household Income	Less than $25,000	437 *5.4%*	326 *4.1%*	849 *10.6%*	167 *2.1%*	1,779 *22.2%*
	$25,000 to 49,999	137 *1.7%*	250 *3.1%*	1,129 *14.1%*	260 *3.2%*	1,776 *22.1%*
	$50,000 to 74,999	44 *0.5%*	201 *2.5%*	853 *10.6%*	261 *3.3%*	1,359 *16.9%*
	$75,000 to 99,999	13 *0.2%*	133 *1.7%*	585 *7.3%*	227 *2.8%*	958 *11.9%*
	$100,000 or more	12 *0.1%*	249 *3.1%*	1,267 *15.8%*	624 *7.8%*	2,152 *26.8%*
	Total	643 *8.0%*	1,159 *14.4%*	4,683 *58.4%*	1,539 *19.2%*	8,024 *100.0%*

Table 5.12: Cross-tabulation of Household Vehicle Ownership Category by Driving Limitation Status in the Attitudinal Dataset (Weighted Dataset, N= 8,024)

		Vehicle Category				
		ZVO	LTE	AAE	HTE	Total
Driving Limitations	No	481	1,051	4,339	1,482	7,353
		6.0%	*13.1%*	*54.1%*	*18.5%*	*91.6%*
	Yes	162	107	344	58	671
		2.0%	*1.3%*	*4.3%*	*.7%*	*8.4%*
	Total	643	1,158	4,683	1,540	8,024
		8.0%	*14.4%*	*58.4%*	*19.2%*	*100.0%*

Table 5.13: Classification of Cases by Household Income Level and Driving Limitations Status in the Attitudinal Dataset (Weighted Dataset, N= 8,024)

			Driving Limitations		Total
			No	Yes	
Household Income	Lower	Number of Cases	3,102	496	3,598
		% within Income Category	*86.2%*	*13.8%*	*100.0%*
		% within Driving Limitation Category	*42.2%*	*73.9%*	*44.8%*
		% of Total	38.7%	6.2%	44.8%
	Higher	Number of Cases	**4,251**	175	4,426
		% within Income Category	***96.0%***	*4.0%*	*100.0%*
		% within Driving Limitation Category	***57.8%***	*26.1%*	*55.2%*
		% of Total	**53.0%**	2.2%	55.2%
Total		Number of Cases	7,353	671	8,024
		% within Income Category	*91.6%*	*8.4%*	*100.0%*
		% within Driving Limitation Category	*100.0%*	*100.0%*	*100.0%*
		% of Total	91.6%	8.4%	100.0%

The highlighted cell in Table 5.13 identifies the target group of respondents who do not have driving limitations and live in a household with median-or-higher annual income. They will be the main object of the analysis in the remainder of this section. Interestingly, the share of total cases falling into this cell (53.0%) is very similar to the comparable share for the NHTS sample (Table 5.5, 53.7%). This increases our confidence that the two weighted samples are comparable on the variables of interest and thus that findings from one are reasonably transferable to the other. Also, please note that for all analyses developed with the attitudinal dataset, the definition of driving limitations is slightly different from that one used for the NHTS dataset. While for the NHTS data we tracked whether the households included *any adult members* with a medical condition precluding driving, the attitudinal dataset includes information on driving limitations only for *the individual completing the survey*. For additional information about the variables included in the attitudinal dataset used in this part of the project, please refer to the description in Section 4.

To investigate the reasons behind the decisions related to vehicle ownership, we analyze the attitudinal profiles of the individuals contained in this target group. In particular, in the remainder of this section, among those in this group, we will compare the attitudinal profiles of the individuals who are in the lower vehicle ownership categories (i.e., individuals that live in ZVO or LTE households) to those in the higher vehicle ownership categories (i.e., individuals that live in AAE or HTE households).

Table 5.14 summarizes the content of the attitudinal variables that are analyzed in the following set of figures. For each variable, a higher value is hypothesized to positively influence an individual('s household) to own zero or fewer-than-expected vehicles. Accordingly, in Figures 5.1 to 5.4, we report the percentage of respondents in that group who have a value of that variable that is

a) higher than the median, for factor scores based on the factor analysis of multiple attitudinal statements; or
b) positive (i.e., either "Agree" or "Strongly Agree"), for attitudinal statements that are measured on a 5-level Likert scale.

In addition, for each attitudinal variable included in the figures, we report in parentheses the mean value of the attitudinal variable for the individuals that have higher-than-the-median/positive attitudes, as described above.

Because it is difficult to mentally process the interaction of five variables at once, we examine the attitudes three at a time, in two sets of combinations: pro-environment, anti-driving, and pro-density (Figures 5.1 and 5.2); and pro-density, pro-transit, and bike-and-walk-liking (Figures 5.3 and 5.4). Using the information for the variables listed in Table 5.14, we created a number of Venn Diagrams that represent the proportion of respondents (respectively in the lower vs. higher vehicle ownership categories) that have high values for each specific combination of three attitudinal variables.

Table 5.14: Definition of the Variables Used for Analysis of Attitudes of High-Income, No Driving Limitations Respondents

Variable Name	Type of Variable	Meaning
pro-environment	Factor score extracted from several attitudinal statements (the highest loading items include the statements "To improve air quality, I am willing to pay a little more to use an electric or other clean-fuel vehicle" and "We should raise price of gasoline to reduce congestion and air pollution")	The respondent has a pro-environmental attitude
anti-driving	Factor score extracted from several attitudinal statements (the highest loading items include the statements "I am willing to reduce driving to improve transportation and air quality" and "I would usually rather have someone else do the driving")	The respondent has an anti-driving attitude
pro-density	Factor score extracted from several attitudinal statements (the highest loading items include the statements "Having shops and services within walking distance of my home is important to me" and "I need to have space between me and my neighbors", the latter having a negative loading on this factor)	The respondent prefers living in a denser, more urban, neighborhood
pro-transit	Factor score extracted from several attitudinal statements (the highest loading items include the statements "I like taking transit" and "I prefer to take transit rather than drive whenever possible")	The respondent likes using public transportation
like bike and walk	Agreement with single attitudinal statement	The respondent likes walking and biking

Figure 5.1 summarizes the attitudinal profiles, for the "pro-environment", "anti-driving", and "pro-density" attitudes, for the individuals who have higher income, no driving limitations and who live in either a ZVO or LTE vehicle-owning household (the numbers in the figure are the percentages of cases from this group that have values for each respective attitudinal variable, or combination of variables, above the median).[10] The comparison with Figure 5.2, which exhibits the profiles for

[10] Percentages for the ZVO/LTE group are often referring to a small number of cases in the dataset. Therefore, care should be used in drawing conclusions on the magnitude of some of these differences.

cases in the target group who are in either the AAE or the HTE categories, allows drawing some interesting conclusions about the reasons for individuals in these groups to be in the lower vehicle ownership categories.

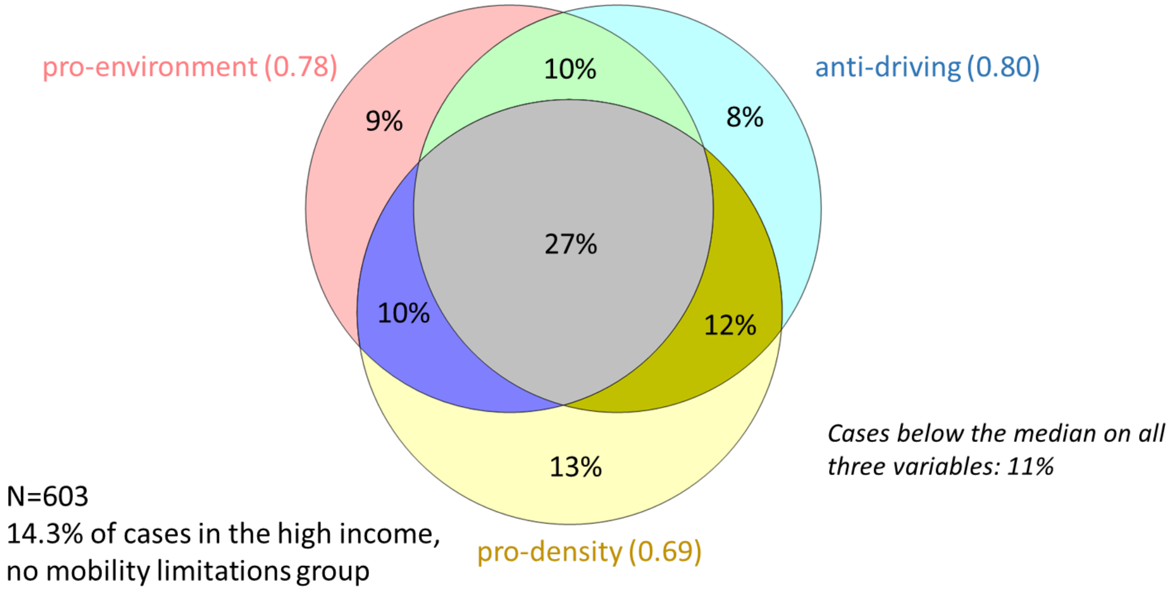
The first observation is that the vast majority of cases in both figures have above-median responses on at least one of the three attitudes, but the ZVO/LTE group has 6 percentage points more: 89% versus 83% for the AAE/HTE group. So having at least one favorable attitude offers a partial, but limited, explanation for the choice to have zero or fewer than expected vehicles. Taken one at a time, individuals that are in the lower vehicle ownership categories are found to have more positive values for all three attitudes that are analyzed in these two figures. For example, 62% of the respondents in the lower vehicle ownership categories (Figure 5.1) are found to have attitudes higher than the median on the pro-density dimension, compared to only 49% of the respondents in the higher vehicle ownership categories (Figure 5.2). Similarly, the shares for the pro-environment dimension are 56% versus 49%, and for the anti-driving dimension are 57% versus 47%.

Second, as shown by the greater amount of overlap of the circles in Figure 5.1 compared to 5.2, the ZVO/LTE group more often has above-median values on *combinations* of these three attitudes. For example, 27% of the target group individuals that are in the lower vehicle ownership categories are found to have positive attitudes for all three dimensions that are analyzed, compared to only 17% of the cases for the counterpart group in the AAE and HTE categories. In all, 41% of the ZVO/LTE target group cases have above-median values for at most one of the three attitudes, whereas 54% of the AAE/HTE cases do. This sizable difference suggests that it may take the congruence of multiple attitudes to effect voluntary reductions in vehicle ownership; having only one supportive attitude may not suffice.

Third, not only is the number of individuals having higher-than-median attitudes larger in the lower than in the higher vehicle ownership categories, but also among these individuals with higher-than-median values, the mean for each attitudinal variable (reported in parentheses after each variable, in Figures 5.1-5.4) is always notably higher for the ZVO/LTE cases than for the AAE/HTE cases.

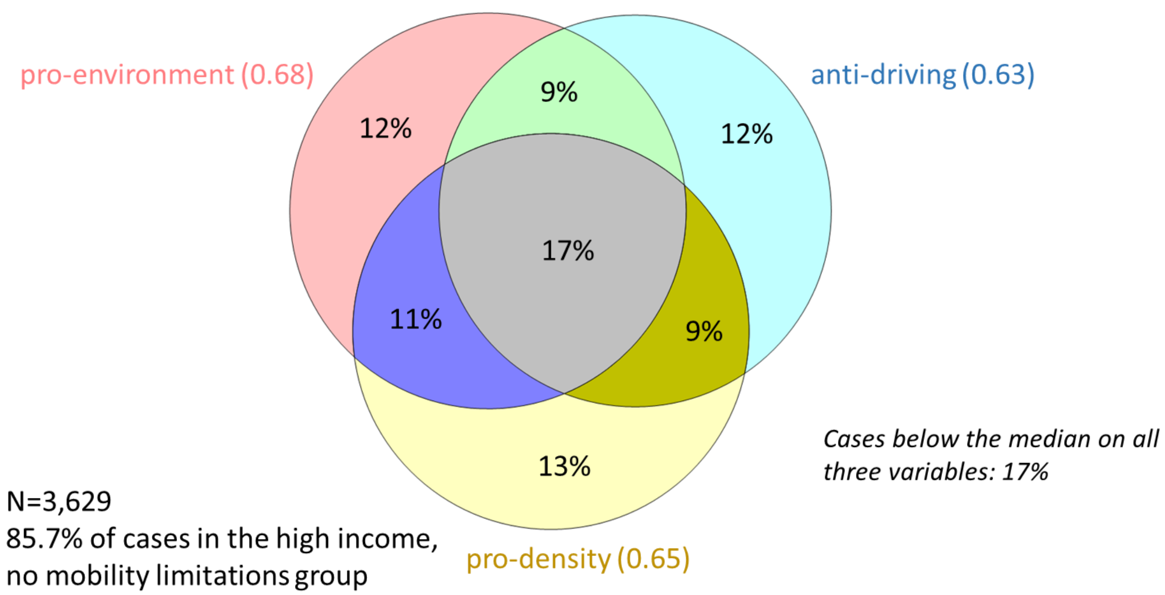
In sum, compared to those with the expected number of vehicles or more, those who have fewer vehicles than expected (1) are *more likely to have attitudes supportive* of a voluntary lower carbon footprint; (2) tend to have *more such attitudes* in combination; and (3) tend to hold those attitudes *more strongly*.

Figure 5.1: Distribution of Pro-environment, Anti-driving, and Pro-density Attitudes among Target-group Individuals who Own Zero- or Fewer-than-expected Vehicles



Note: The values in parentheses represent the mean for each attitudinal variable measured for the individuals that have higher-than-the-median attitude for that dimension. These attitudes are standardized to have mean 0 and variance 1, so a mean close to 1 would considered relatively strong in this context.

Figure 5.2: Distribution of Pro-environment, Anti-driving, and Pro-density Attitudes among Target-group Individuals who Own About-as-expected or More-than-expected Vehicles



Note: The values in parentheses represent the mean for each attitudinal variable measured for the individuals that have higher-than-the-median attitude for that dimension. These attitudes are standardized to have mean 0 and variance 1, so a mean close to 1 would considered relatively strong in this context.

Similarly to what was done for the first group of attitudinal variables, Figures 5.3 and 5.4 summarize the distribution of the pro-density, pro-transit, and bike-and-walk-liking attitudes of individuals who do not have driving limitations and live in higher-income households, who own zero or fewer-than-expected vehicles, or about-as-expected or more-than-expected vehicles, respectively.

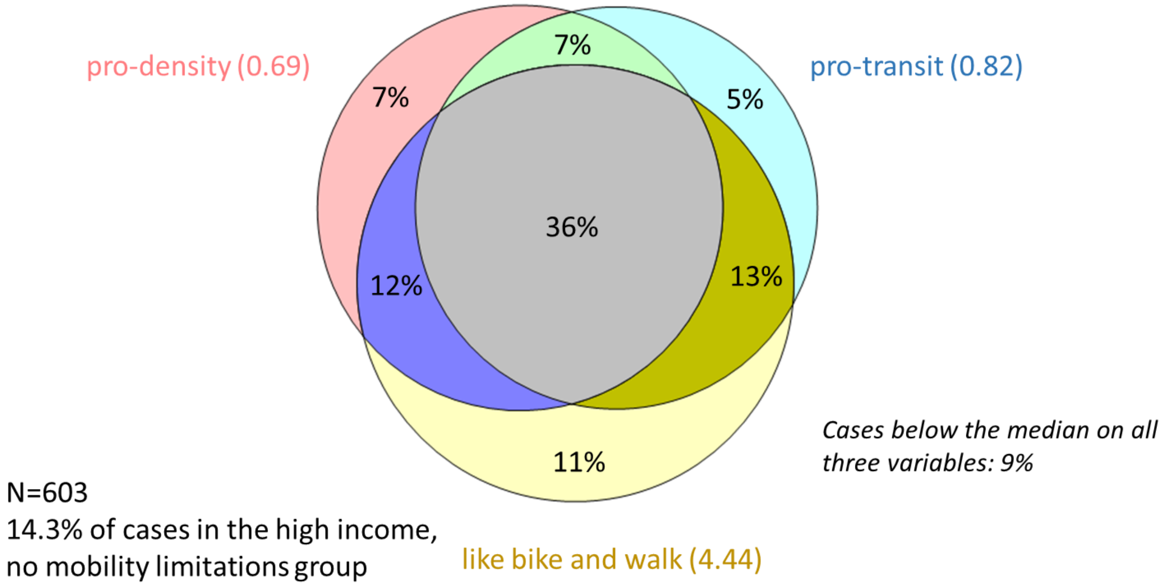
Also for this second group of attitudinal variables, individuals that live in ZVO or LTE households tend to have attitudinal profiles that differ in several ways from those of the individuals from AAE and HTE households. First, the vast majority of individuals who live in lower vehicle ownership category households have above-the-median responses also for all attitudinal variables in this second group. For example, 61% of individuals in ZVO or LTE households are found to have attitudes higher than the median on the pro-transit dimension, compared to only 48% of the individuals in AAE and HTE households.

In addition, similar to what was observed for the first group of attitudinal variables, ZVO/LTE individuals more often have above-median values for the combination of all three attitudes, as well as for all pairs of two attitudes. For example, 36% of the target group individuals that are in the lower vehicle ownership categories are found to have positive attitudes for all three dimensions pro-density, pro-transit and bike/walk-liking, compared to only 23% of the cases for the counterpart group in the AAE and HTE categories. In all, 31% of the ZVO/LTE target group cases have above-median values for at most one of this second group of three attitudes, whereas 46% of the AAE/HTE cases do.

In particular, it is interesting to note that a sizable share (64%) of individuals in the AAE/HTE group has a higher-than-median attitude on the "like bike and walk" dimension, and in that group, the mean attitude (4.35) is not much lower than that for the higher-than-median cases in the ZVO/LTE group (4.44). This finding suggests that the action of only this attitude is not enough to motivate voluntary reductions in vehicle ownership, consistent with some other studies in which bicycling and walking have been found to partially augment rather than replace vehicle trips (Handy and Clifton, 2001; Piatkowski et al., 2015).

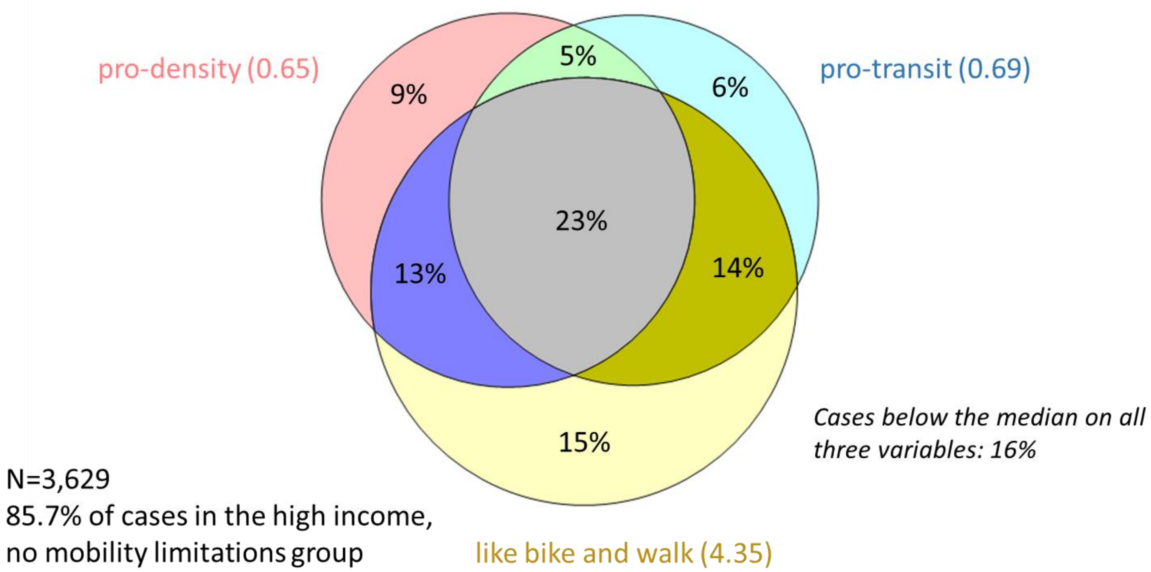
The attitudinal profiles that are discussed in this section provide some information on the reasons behind the decisions regarding vehicle ownership of the individuals who do not have income or driving limitations, and who therefore have more space to make voluntary choices regarding vehicle ownership and their desired mobility patterns. The analyzed attitudes certainly account for *some part* of the observed differences in vehicle ownership. However, the attitudinal patterns do not seem extremely different across groups. This might be a reason for their rather modest contribution to the explanation of vehicle ownership choices in the models that were estimated for Task A.3. Still, the attitudinal profiles of the members of a household do have an effect on mobility patterns and vehicle ownership levels, the more so when multiple conducive attitudes are held simultaneously (as shown in the figures in this section). This suggests that to increase the voluntary choice to reduce vehicle ownership, there is value in trying to influence individuals' attitudes, preferably in combination.

Figure 5.3: Distribution of Pro-density, Pro-transit, and Bike-and-walk-liking Attitudes among Target-group Individuals who Own Zero- or Fewer-than-expected Vehicles



Note: The values in parentheses represent the mean for each attitudinal variable measured for the individuals that have higher-than-the-median attitude for that dimension. The pro-density and pro-transit attitudes are standardized to have mean 0 and variance 1, whereas the like bike/walk attitude is a single item measured on a 5-point scale.

Figure 5.4: Distribution of Pro-density, Pro-transit, and Bike-and-walk-liking Attitudes among Target-group Individuals who Own About-as-expected or More-than-expected Vehicles



Note: The values in parentheses represent the mean for each attitudinal variable measured for the individuals that have higher-than-the-median attitude for that dimension. The pro-density and pro-transit attitudes are standardized to have mean 0 and variance 1, whereas the like bike/walk attitude is a single item measured on a 5-point scale.

It is important to remember that several limitations apply to this analysis, for example due to the lack of comprehensive information about the residential location and the characteristics of the built environment in the attitudinal dataset. This limits the ability to ascertain the impact of individual attitudes versus the impact of geography and residential location on vehicle ownership decisions. Future extensions of this research should focus on analyzing these relationships, in particular through the collection of a comprehensive behavioral and attitudinal dataset, with complete information on an individual's travel behavior, vehicle ownership, attitudinal profiles, individual background and residential location; and the estimation of models that control for attitudes in general, and residential self-selection in particular.

6. Task A.5: The Impact of Geographic Location and Urban Form on Vehicle Ownership and VMT

The earlier tasks examined the impacts of socioeconomic traits and attitudes on vehicle ownership and VMT. However, understanding the role of land use characteristics in affecting households' vehicle ownership and VMT is also important, for many reasons. Clearly, we expect the characteristics of the residential neighborhood where a household lives to influence its decision to own fewer or more than the expected number of vehicles (Cao et al., 2007; Salon et al., 2012): living in a denser urban environment that is well-served by transit and has many activity locations near home makes it more likely that a household can and will forgo a(n additional) car. From a planning and policy-making perspective, understanding how the built environment influences vehicle ownership and VMT is important to support the development (and properly forecast the related effects) of policies targeted at improving sustainability, such as the Sustainable Community Strategies mandated in California by Senate Bill (SB) 375.

To investigate this topic, we classified all residential locations in the NHTS sample into a small number of geographical categories. We then re-estimated the multinomial logit models allowing all coefficients in the vehicle ownership and VMT models to differ by geographic area. Because the analysis was conducted on the entire NHTS sample, but the interest focused on California, the segmentation variables needed to be generic in the definition of the geographic regions and neighborhood types, and not specific to a geographic location such as the San Francisco Bay Area, or Minnesota.

Accordingly, in this part of the study we created six geographic clusters along two dimensions of variation: regional and local (neighborhood). The regional dimension is based on the population size and the presence of rail transit in the metropolitan statistical area (MSA), resulting in three categories: *smaller region* (i.e., total population less than one million), *larger region* (i.e., total population greater than one million) *with no rail*, and *larger region with rail*. The local dimension is based on the computed *density factor score of the household's residential neighborhood*, classified as *lower-density* or *higher-density*, on the basis of a factor score (as explained below) created from four density variables. Crossing these two dimensions created six clusters, where Clusters 1 through 3 represent the lower-density neighborhoods (LDNs), in increasing order of regional importance, and Clusters 4 through 6 represent the higher-density neighborhoods (HDNs), in the same regional order. Figure 6.1 portrays the six clusters that were identified in this study, together with the number of cases that are included in each cluster in the weighted NHTS dataset.

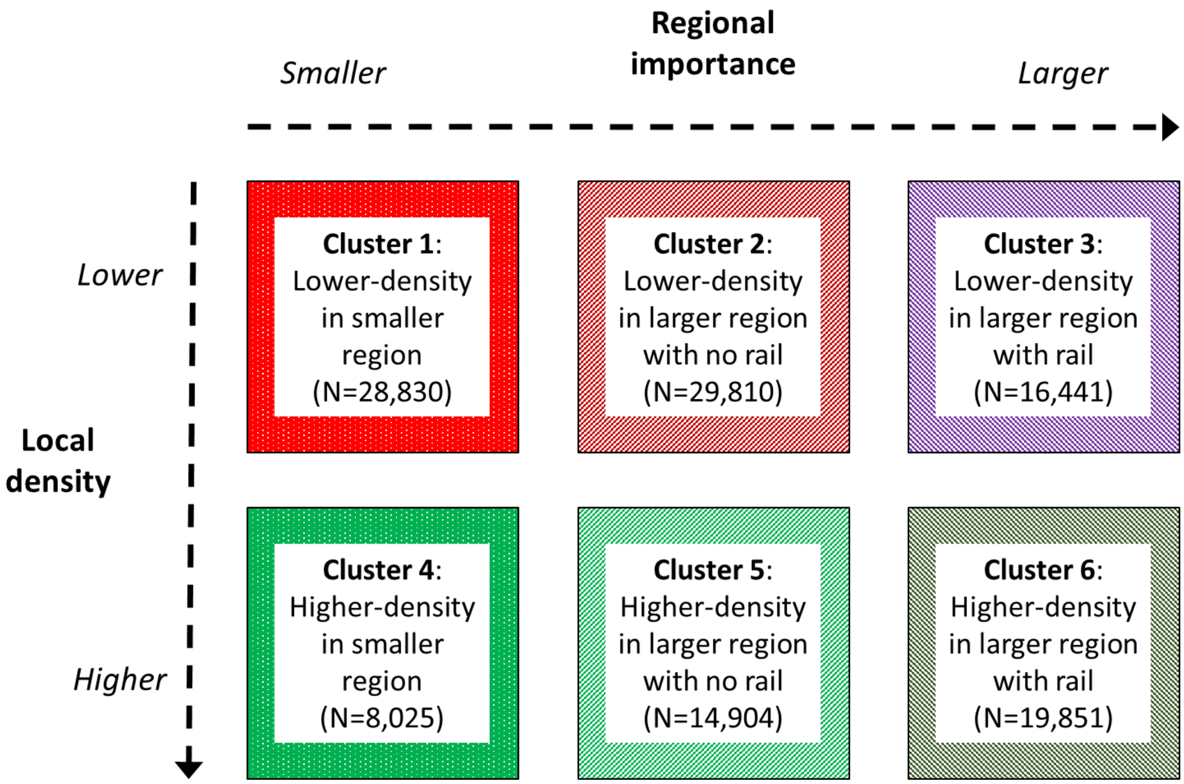


Figure 6.1: Identification of Six Clusters of Neighborhood Types by Regional Dimension and Local Residential Density, and Sample Sizes in the Weighted NHTS Dataset

We measured neighborhood density through a density factor score, which was created using principal components analysis on four of seven different built environment variables available in the NHTS dataset. These four variables include (1) housing units per square mile, (2) percentage of housing units that are renter-occupied, (3) population per square mile, and (4) employees per square mile living in the census tract. This computed density variable score was then used in the models of vehicle ownership category and VMT that were estimated either with the pooled sample (containing all cases in the NHTS dataset) or through segmenting the sample by the six geographic clusters described above. A household's residential neighborhood was classified as lower-density (Clusters 1 through 3) or higher-density (Clusters 4 through 6) when the values of the residential density factor score for that household was respectively below or above the sample mean[11].

[11] The factor score is a weighted linear combination of the four variables on which it is based, where the weights are proportional to the correlation of the respective variable to the underlying (unobserved) factor which is approximated by the computed score. Factor scores are typically standardized, and thus have mean 0 and variance 1, generally ranging between -3 and 3 in value. A score of -3 would refer to a neighborhood with low values on the four constituent variables, while a score of +3 would indicate a neighborhood with high values (very dense).

6.1 Impact of Geographic Location and Urban Form on Vehicle Ownership

Geographical elements such as presence of rail and density variables were incorporated into the models estimated using the NHTS data, to account for the influence of land use characteristics on vehicle ownership. Not surprisingly, accounting for land use features such as density and the characteristics of the urban region improves the ability to predict the household vehicle ownership category by a non-trivial amount.

Table 6.1: Summary of the Goodness of Fit of the Vehicle Ownership Category Models (NHTS Dataset, Weighted Sample)

Model Specification	N (Sample Size)	ρ^2 (EL Base)	Percentage of Cases Correctly Classified	% Improvement of ρ^2 (EL base) over Base Model	% Improvement in Cases Correctly Classified over Base Model	Percentage-Point Improvement in Cases Correctly Classified
ASCs only (Base Model)	130,477	0.2398	44.7%	-	-	-
ASCs + Income	130,472	0.2787	45.7%	*16.2%*	*2.24%*	1.0
ASCs + Income + Driving Limitations	130,329	0.2857	46.2%	*19.1%*	*3.36%*	1.5
ASCs + Income + Driving Limitations + Density (Pooled Sample)	130,329	0.3211	47.8%	*33.9%*	*6.94%*	3.1
ASCs + Income + Driving Limitations + Density (Joint Segmented Model)	130,329	0.3264	48.3%	*36.1%*	*8.05%*	3.6

We included the density factor score as another explanatory variable into the model containing income and driving limitations developed in Task A.2. As expected, the characteristics of the built environment are an important predictor of vehicle ownership: just adding density to the best model estimated on the entire (pooled) sample from Task A.2 (income + driving limitation) improves the EL-base ρ^2 from 0.2857 to 0.3211 (a 12.4% increase), as seen in Table 6.1.

Table 6.2: Including Density in the Best Vehicle Ownership Category Multinomial Logit Model (NHTS Dataset, Weighted N = 130,303)

	ZVO	LTE	AAE	HTE
Income Level 1 (Less than $25,000)	**2.823** (58.91)	**0.441** (16.56)	(base)	**-0.875** (-35.17)
Income Level 2 ($25,000 to 49,999)	**1.486** (29.14)	**0.319** (12.17)	(base)	**-0.505** (-23.93)
Income Level 3 ($50,000 to 74,999)	**0.475** (7.56)	**0.127** (4.35)	(base)	**-0.201** (-9.52)
Income Level 4 ($75,000 to 99,999)	**0.302** (4.09)	**0.107** (3.24)	(base)	**-0.114** (-4.88)
Income Level 5 ($100,000 or larger)	(base)	(base)	(base)	(base)
Driving Limitation (DV)	**1.546** (50.31)	**0.087*** (2.39)	(base)	
Density Factor Score	**0.932** (91.75)	**0.389** (46.20)	(base)	**-0.365** (-35.24)
Constant	**-4.314** (-92.82)	**-1.959** (-102.77)	(base)	**-1.016** (-77.89)
Log likelihood final model	-122627.53			
Log likelihood constants only	-137063.64			
Log-likelihood equally-likely model	-180638.23			
ρ^2 (equally-likely base)	0.3211			
Adjusted ρ^2 squared (EL base)	0.3210			

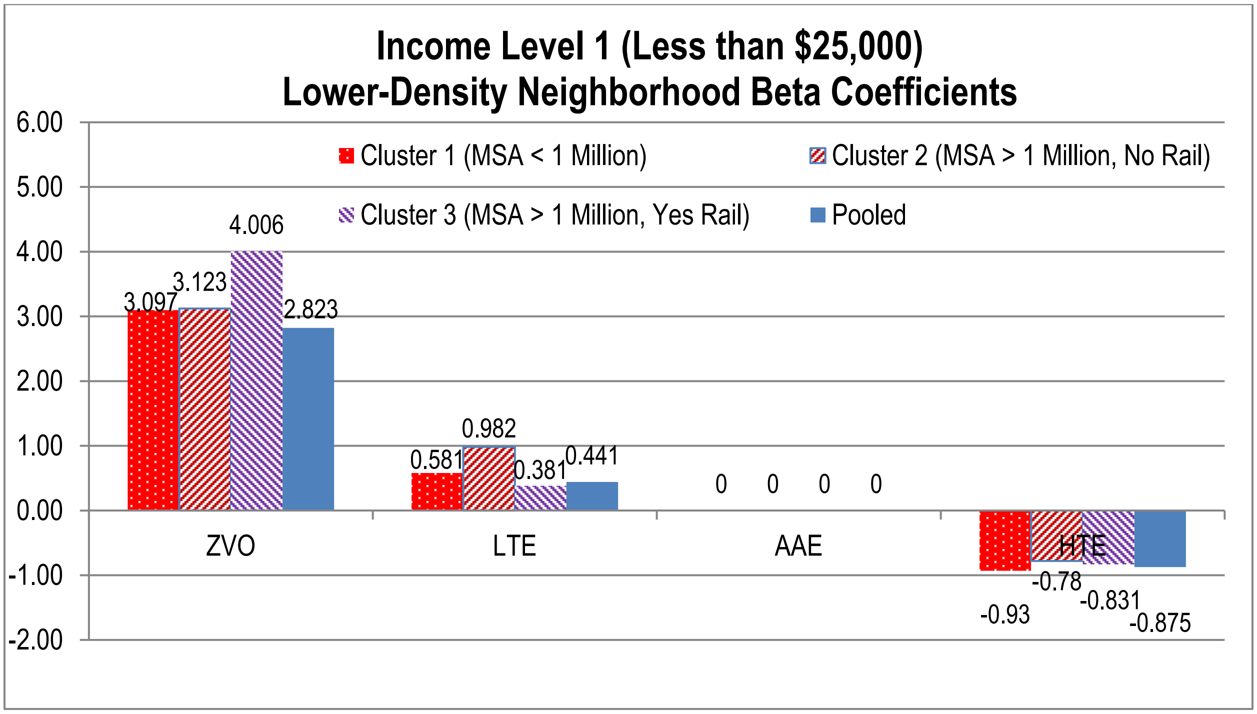
*Note: Coefficients are in bold and t-statistics are reported in parentheses. All estimated coefficients are significant at least at the 0.1% level, unless otherwise noted. Coefficients marked with * are significant at the 5% level.*

All else equal, and as expected, the higher the density in the neighborhood where a household lives (as captured by the density factor score), the higher the probability that the household is in the ZVO or LTE categories. Thus, households that live in higher-density neighborhoods are significantly more likely to own zero or fewer than expected vehicles (i.e., be in the ZVO or LTE categories), compared to households that live in lower-density neighborhoods.

We accounted for the role of location characteristics not only by incorporating density as an explanatory variable, but also by allowing the impacts of all the *other* variables in the model to differ by location type. To do this, we segmented the NHTS sample into six categories defined by local and regional characteristics - namely (1) the density of the household's residential neighborhood ("higher" than the mean value of the factor score versus "lower" than the mean), and (2) the region size and presence of rail (with three "regional status" categories: Metropolitan

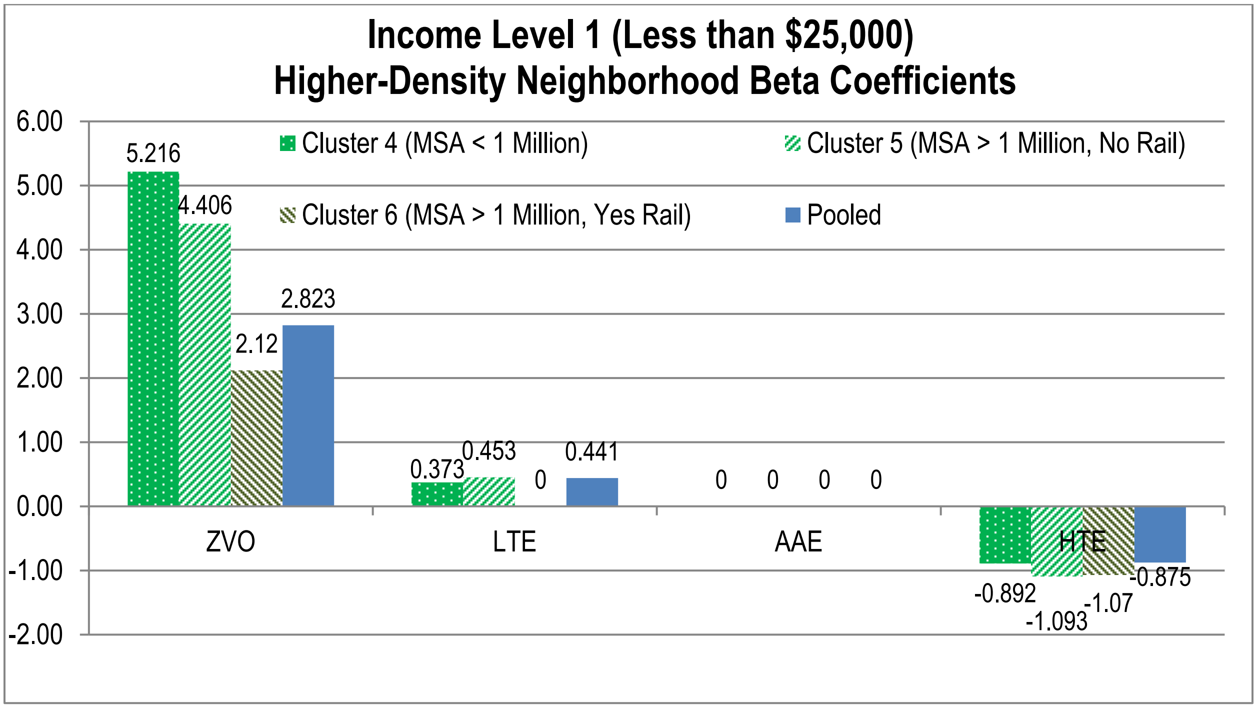
Statistical Areas (MSAs) with population smaller than one million, and MSAs with population greater than one million without and with rail mass transit) - and estimated different models for each segment. Thus, we re-estimated the multinomial logit models with the same specifications mentioned above (i.e., including residential density) for each of the six clusters individually. The EL-base ρ^2 values for these six models range from 0.2724 to 0.3722, with an overall ρ^2 of 0.3264 for the joint segmented model. The share of correctly classified cases for these six clusters ranges from 44.8% to 51.3%, which is rather large improvement (6.6 percentage points, p.p., or 14.5%) compared to all previous models that did not include the density factor score. Overall, the joint segmented model is capable of correctly classifying 48.3% of all the cases, which is 0.5 p.p. (1.0%) better than the pooled dataset's percentage of cases correctly classified.

Figure 6.2: Estimated Coefficients for Income Level 1 in Lower-density Neighborhoods



The analysis of the segmented models by neighborhood type allows us to investigate the different effects of the explanatory variables across different land use types. For example, how does the influence on vehicle-ownership category of being in a low-income household differ by residential neighborhood type? Figures 6.2 to 6.9 visually compare the estimated beta coefficients for the four income categories (using the highest income category, $100,000 or more, as reference) in lower-density neighborhoods (LDNs) vs. higher-density neighborhoods (HDNs), and in different region size/rail categories, as defined by the six clusters. The coefficients for the AAE vehicle ownership category are always equal to zero (and included in the figures for comparison), as AAE was the base alternative in the estimation of the multinomial logit models. Each figure also includes the coefficients for the pooled (across all six neighborhood types) model, for comparison.

Figure 6.3: Estimated Coefficients for Income Level 1 in Higher-density Neighborhoods



Looking at Figures 6.2 and 6.3, as expected, the beta coefficients for the lowest income category decrease across the four vehicle ownership categories going from left to right, with households that are in the lowest income category being far more likely to have zero vehicles in the household, and less likely to have more vehicles than expected. Interestingly, however, the effects of belonging to this income category – particularly on the utility of owning zero vehicles – vary significantly by neighborhood type. Specifically, in *higher-density* neighborhoods (Figure 6.3), the income level 1 coefficient for the ZVO category *decreases* with the region's size/rail status, meaning (conversely) that the (positive) impact of having low income on owning zero vehicles is strongest in smaller regions, with fewer than 1 million inhabitants. By contrast, in *lower-density* neighborhoods (Figure 6.2), the coefficient *increases* with regional status, meaning that the impact of having low income on owning zero vehicles is strongest in larger regions with rail.

Figure 6.4: Estimated Coefficients for Income Level 2 in Lower-density Neighborhoods

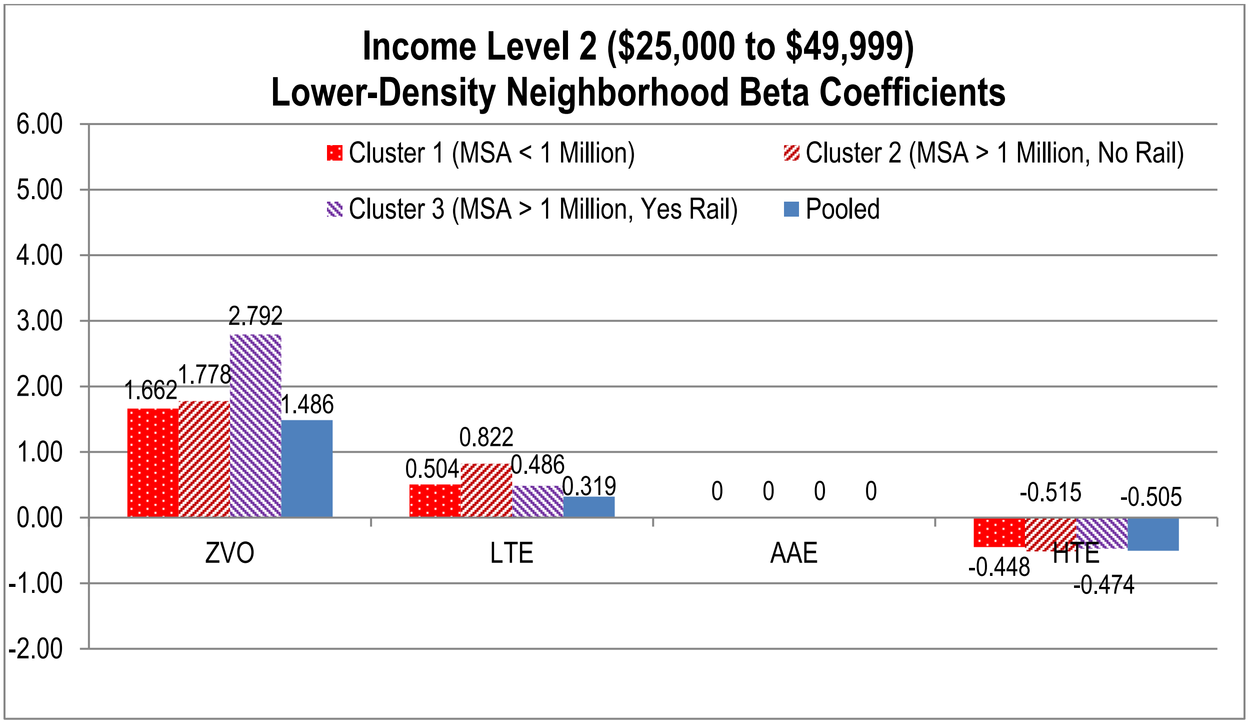


Figure 6.5: Estimated Coefficients for Income Level 2 in Higher-density Neighborhoods

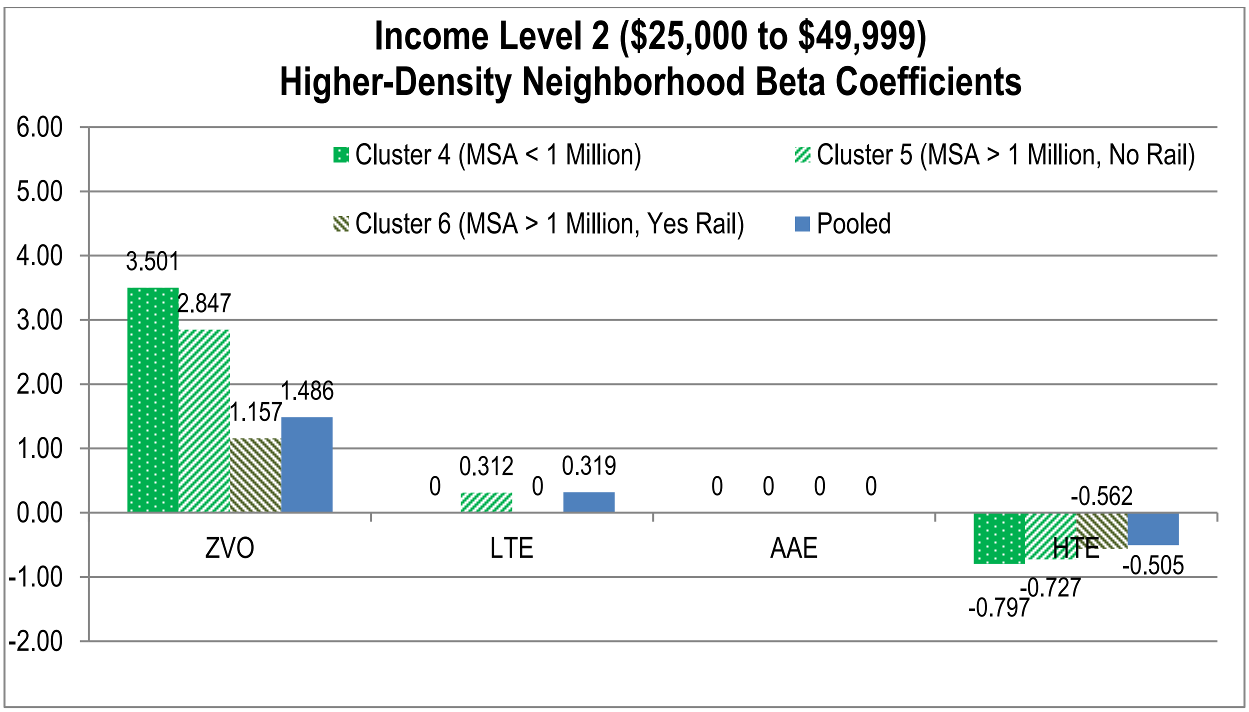


Figure 6.6: Estimated Coefficients for Income Level 3 in Lower-density Neighborhoods

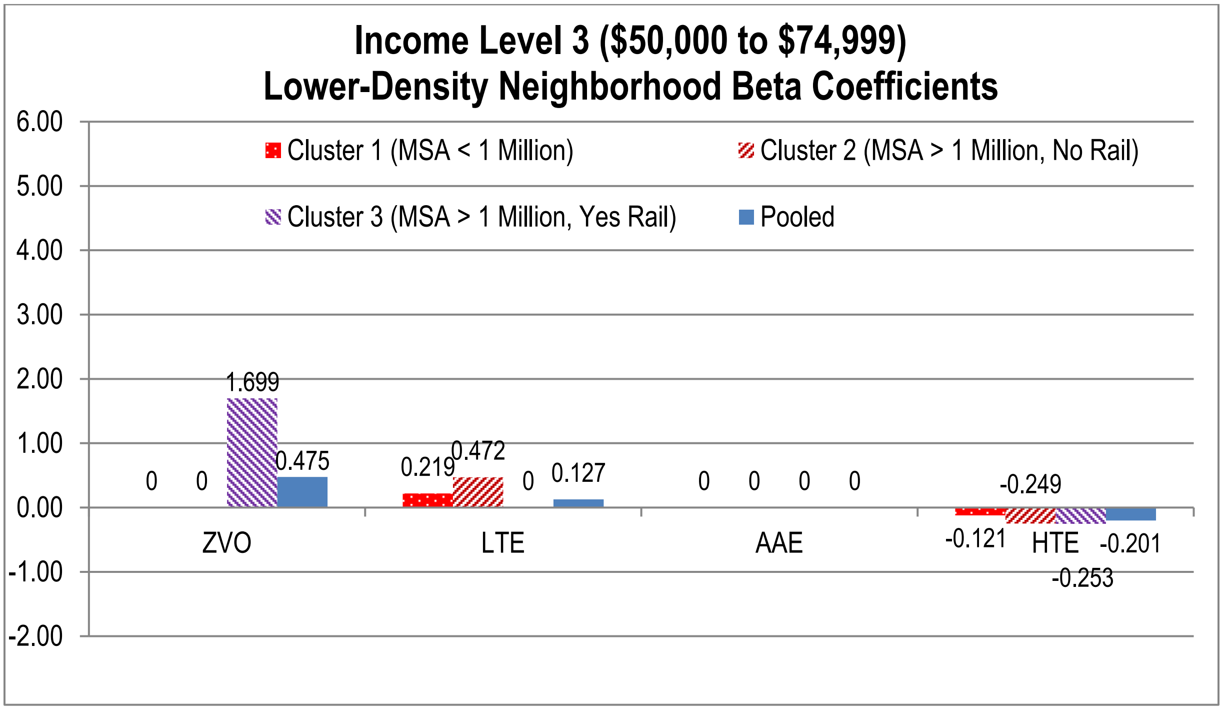
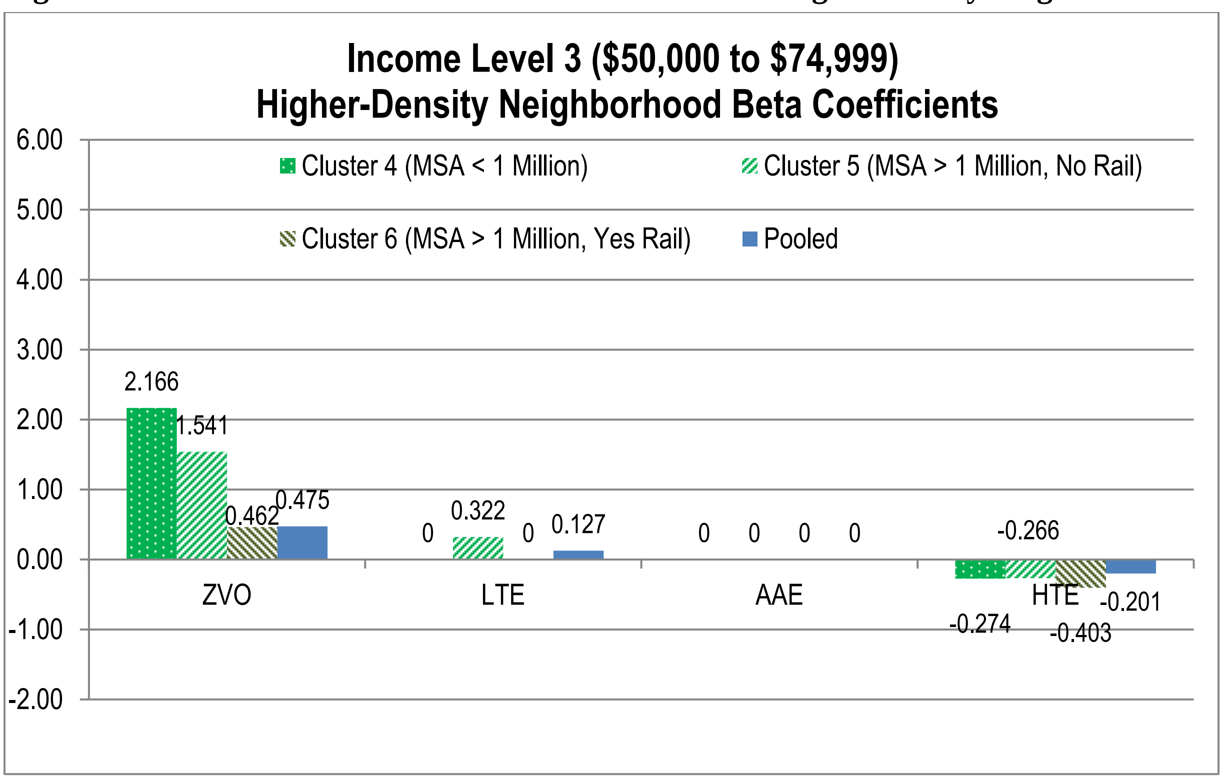


Figure 6.7: Estimated Coefficients for Income Level 3 in Higher-density Neighborhoods



The effects of low income on the likelihood to be in the LTE vehicle owning category are similar across density categories, but are larger in magnitude in lower-density neighborhoods. In both cases, interestingly, the extent to which low-income households are more likely than others to have LTE vehicle ownership is higher for larger regions without rail than for either of the other two types of regions. This pattern likely represents the combined effects of income and land use characteristics on vehicle ownership: low income supposedly acts as a limitation on the vehicle ownership of a household to a larger degree in those areas where transportation alternatives are less available (in the other areas, reductions in vehicle ownership are expected to be driven by voluntary reductions in vehicle ownership associated with the effects of the increased availability of transportation options). Similar, but smaller-magnitude, effects are also associated with the effects of a household belonging to the second lowest income category 2 (Figures 6.4 and 6.5).

The comparisons among Figures 6.6 and 6.7 (Income Level 3) and Figures 6.8 and 6.9 (Income Level 4) highlight similar patterns. Also in these figures, as in all other figures in this section, the coefficients for the AAE vehicle ownership category are included for comparison, but they are always equal to zero as AAE was the base alternative in the estimation of the multinomial logit models. Each figure also includes the coefficients for the pooled (across all six neighborhood types) model, for comparison. Note that the value of some coefficients for the ZVO and LTE groups in these figures are reported as zero, as those coefficients were not found to be statistically significant in the estimated models.

Not surprisingly, the effects of income (relative to the highest-income category) on the likelihood of owning zero or LTE vehicles tend to decrease as income increases, with the effects of income level 4 virtually indistinguishable from those of the base income level 5 for LDNs. Still, an interesting effect is observed in particular for households in the medium/medium-high income categories (3 and 4), who are more likely not to own a vehicle (compared to the highest income category, used as base category) if they live in higher-density neighborhoods, and therefore in neighborhoods that provide more opportunities to travel with non-auto modes. This effect is stronger in smaller regions (with fewer than 1 million inhabitants).

A more straightforward interpretation of these results overall is as follows. As income increases, households become more and more similar to the highest-income (reference) households in their propensity to own vehicles or not. This is not surprising, of course, but the interesting part is that convergence to the reference category is faster for LDNs, especially for regions (of any size) without rail, than for HDNs. Overall, households in the highest-income category are more likely than others to own vehicles no matter in which of the six area types they live in. However, the models indicate that for HDN households, this principle especially holds true for smaller regions, because for HDNs in larger regions even the wealthy may be more inclined to meet their travel needs without a car. For LDN households, the principle applies most strongly to large regions with rail, because for LDNs in other cities even lower-income households are more likely to need a car, and thus to have a propensity to own vehicles that is closer to that of highest-income households.

Figure 6.8: Estimated Coefficients for Income Level 4 in Lower-density Neighborhoods

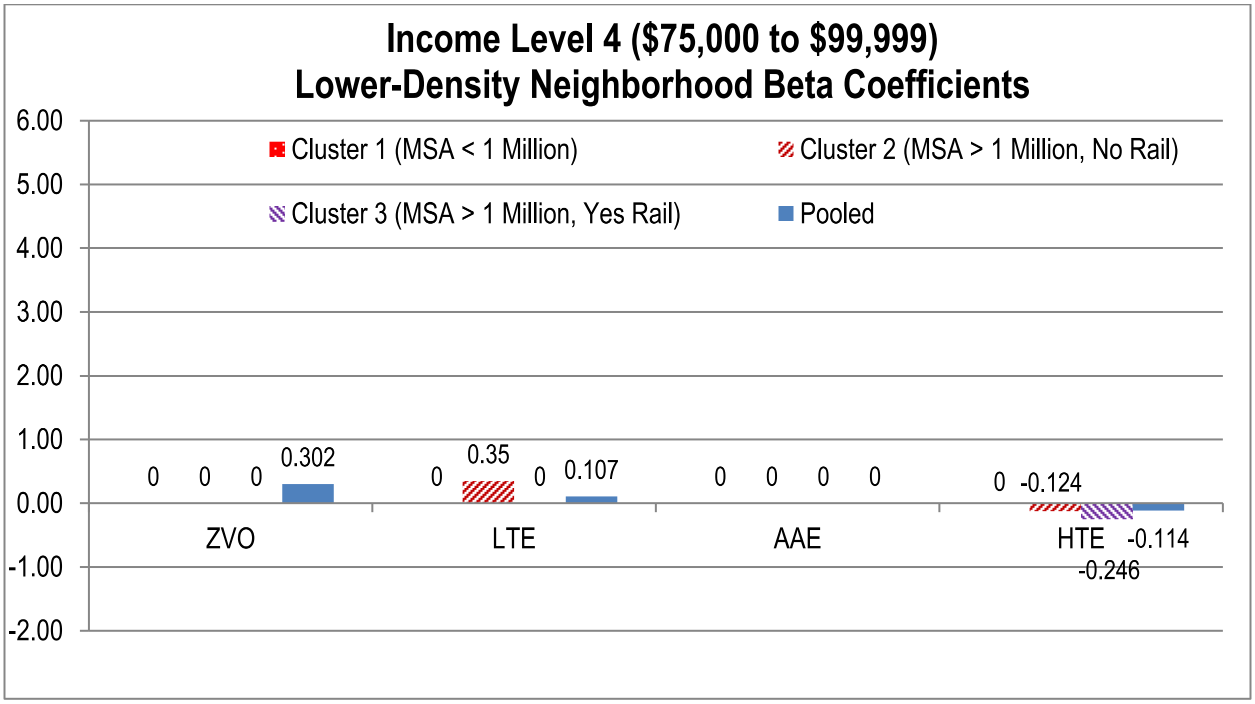
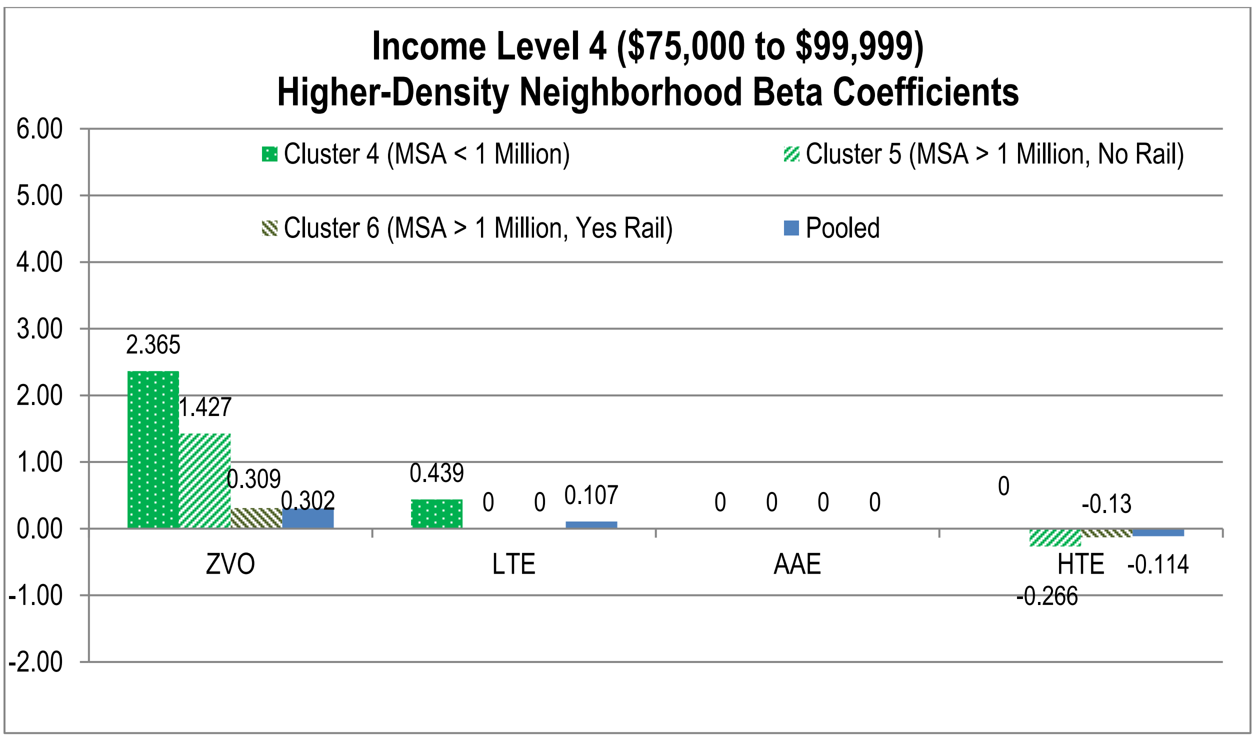


Figure 6.9: Estimated Coefficients for Income Level 4 in Higher-density Neighborhoods



Thus, we find that the convergence between wealthy and less-wealthy households, illustrated by the coefficient bars in the figures getting closer to zero, is from different directions depending on the interaction between region type and neighborhood density. As income increases, households become more and more similar to the highest-income households in their propensity to own vehicles or not. In lower-density neighborhoods, as regional status *diminishes* (moving from large metropolitan areas served by rail to smaller cities) the less-wealthy become similarly likely (to the wealthy) to own cars (mostly *out of necessity*). In higher-density neighborhoods, by contrast, as regional status *increases* (moving from smaller cities to large metropolitan areas served by rail), the wealthy become similarly likely (to the less-wealthy) to *not* own cars (mostly *out of choice*).

Figures 6.10 and 6.11 compare the estimated coefficients associated with the density factor score for low density clusters vs. high density clusters. Across all clusters, where significant, density has a positive coefficient for the ZVO and LTE categories, meaning that all else equal, increases in density increase the utility of owning zero or fewer-than-expected vehicles (and conversely for HTE).

Interestingly, we again see an interaction between local and regional land-use characteristics, in that households living in LDNs are *less* responsive to increased neighborhood density if they are in large regions with rail compared to the other two regional types, whereas households living in HDNs are *more* responsive to increased neighborhood density if they live in large regions, especially those with rail, compared to those in smaller regions. This differential effect of density in the various combinations of land use settings and neighborhood types is most likely due to the differential "room" available to households for making adjustments in VO, depending on the location in which they live and the available transportation alternatives, and to some "threshold" effects: an increase in urban density leads to different changes in vehicle ownership depending on the initial density level and local characteristics of the neighborhood (a finding of particular importance in planning and policy evaluation).

Figure 6.10: Density Factor Score for Lower-density Neighborhoods

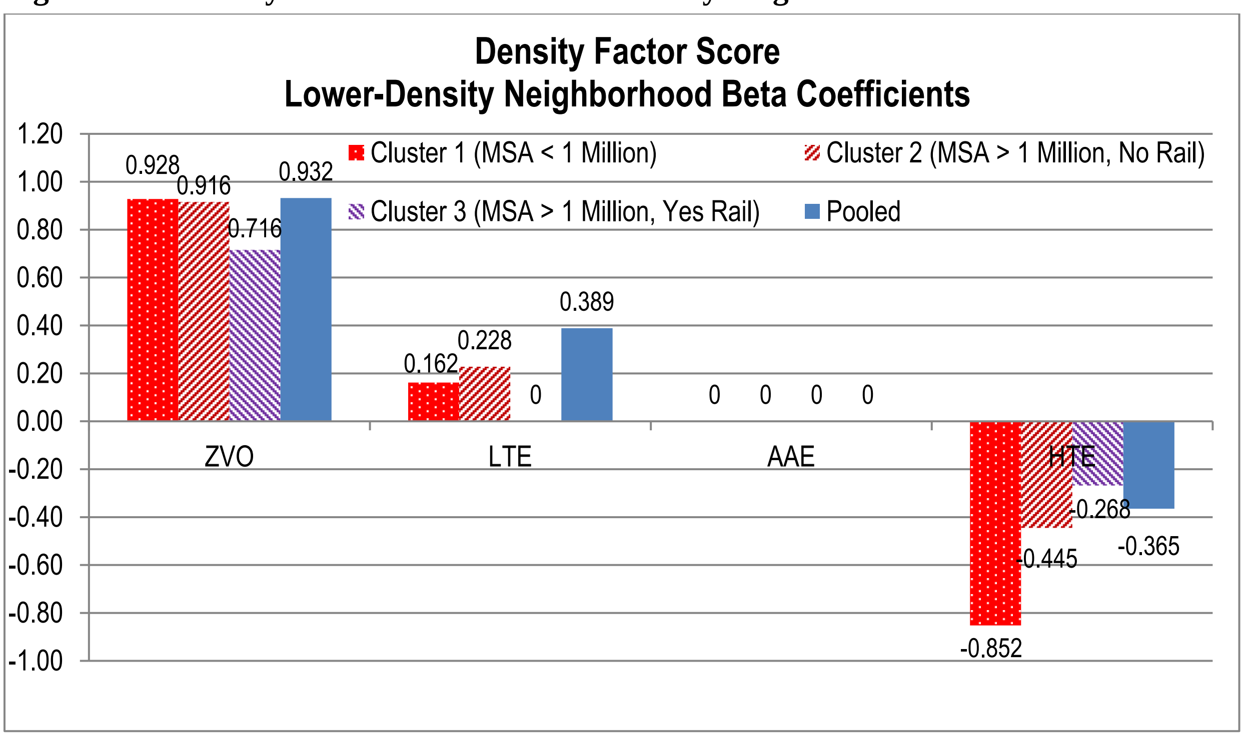
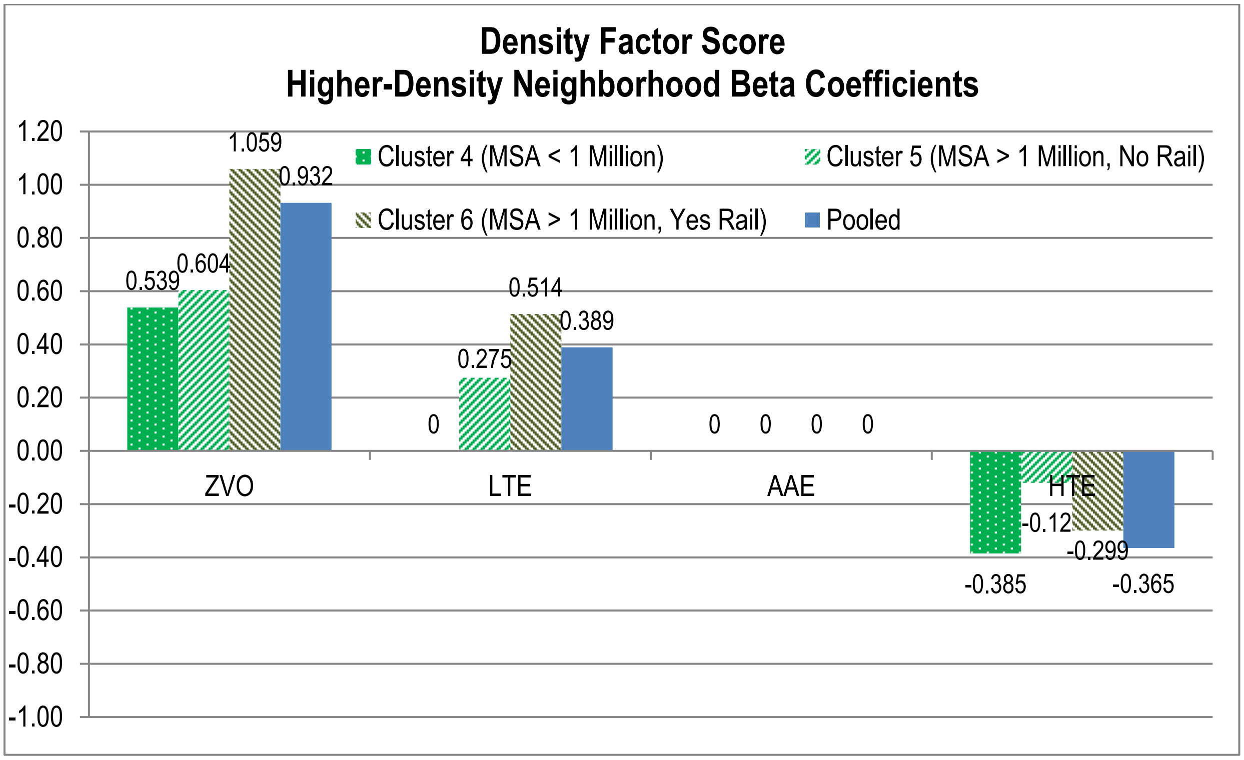


Figure 6.11: Density Factor Score for Higher-density Neighborhoods



6.2 Impact of Geographic Location and Urban Form on Household VMT

After analyzing vehicle ownership models, we estimated the ln(VMT+1) models first for the pooled sample, then for the six different neighborhood types (segmented models). Table 6.3 reports the estimated coefficients in the final segmented models and the pooled model of ln(VMT+1), which explicitly account for the land use characteristics (specifically, density) of the neighborhood where the household lives. In the discussion below, we focus on the residential location relationships to VMT; the sociodemographic relationships are quite standard.

As expected, increases in density are associated with decreases in VMT, across all six neighborhood categories. This association is not uniform across categories, however. For example, we find that urban density has much larger-magnitude effects on VMT, thus predicting larger reductions in VMT when density increases, for households that live in lower-density neighborhoods (Clusters 1-3). The size of the region also counts, with the largest reductions in VMT being associated with an increase in density for the households that live in lower-density neighborhoods in smaller regions, e.g. in a metropolitan statistical area smaller than 1 million (Cluster 1). Specifically among the higher-density neighborhoods (Clusters 4-6), however, the strongest relationship between density and VMT is found in larger regions with rail (Cluster 6): urban density is associated with a larger reduction in VMT for households that live in these areas compared to other areas, after controlling for the impacts of other variables.

Table 6.3: Estimated Coefficients for the Final Models of ln(VMT+1) that Control for Land Use Characteristics (N=119,408, Weighted Dataset)

	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6	Pooled Dataset
Urban Density	-0.320	-0.239	-0.205	-0.139	-0.027	-0.150	-0.155
Driving Limitation	-0.101	-0.151	-0.055	-0.117	-0.127	-0.116	-0.117
Income Level 1 (Less than $25K)	-0.587	-0.606	-0.497	-0.680	-0.529	-0.467	-0.545
Income Level 2 ($25K to 49,999)	-0.326	-0.359	-0.389	-0.394	-0.327	-0.334	-0.329
Income Level 3 ($50K to 74,999)	-0.111	-0.151	-0.112	-0.192	-0.168	-0.138	-0.124
Income Level 4 ($75K to 99,999)	-0.068	-0.071	-0.055	-0.129	-0.079	-0.077	-0.062
Number of Driving Workers	0.583	0.527	0.513	0.606	0.584	0.543	0.553
Number of Driving Non-Workers	0.413	0.346	0.360	0.405	0.383	0.347	0.373
Number of Non-Driving Workers	0.083	0.098	0.137	0.091	0.088	-0.023	0.059
Number of Non-Driving Non-Workers	0.074	0.047	0.004	0.056	0.155	0.127	0.078
Number of Children Under Age 16	0.120	0.113	0.116	0.166	0.144	0.134	0.127
Constant	8.727	8.883	8.859	8.839	8.783	8.861	8.864

Table 6.4 reports the goodness of fit measures for these models: the best model explains some 34% of the variance in ln(VMT+1), which is considered good for disaggregate travel behavior models. Including the density factor score substantially improves the R^2 compared to the models without it. Further, the segmentation of the model, by allowing coefficients to vary for different combinations of neighborhood density levels and region size/availability of rail transit, additionally improves the ability of the model to explain the amount of vehicle travel for households that live in different land use settings. The unconstrained (segmented) model has an R^2 value that is 1.4% better than the constrained (pooled) model.

Table 6.4: Summary of the Goodness of Fit for the ln(VMT+1) Models (NHTS Dataset, Weighted Dataset)

Model Specification	N (Sample Size)	R^2	Adjusted R^2	% Improvement of R^2 over Base Model
Number of Workers, Drivers, and Children (Base Model)	119,413	0.2872	0.2872	-
Number of Workers, Drivers, and Children + Income + Driving Limitation	119,408	0.3204	0.3203	*11.56%*
Number of Workers, Drivers, and Children + Income + Driving Limitation + Density Factor Score (Pooled or Constrained Model)	119,408	0.3382	0.3382	*17.76%*
Number of Workers, Drivers, and Children + Income + Driving Limitation + Density Factor Score (Joint Segmented or Unconstrained Model)	119,408	0.3429	0.3425	*19.39%*

To better understand the influence of land use patterns on VMT, we used the segmented model that was estimated in this part of the project to predict annual VMT for each household in the NHTS dataset, if that household were to live in each of the six different neighborhood types, and then computed averages for each cluster. Specifically, for each household, we used the coefficients of the VMT model for Cluster 1 (multiplying that household's characteristics) to predict that household's VMT if it were to live in Cluster 1, and similarly for the other five clusters. This measure allows estimating the impacts that land use features have on these households' VMT (e.g. the changes in VMT that would be obtained if the households with these characteristics were moving from one neighborhood type to another one, holding everything else, e.g. income and household composition, constant). As expected, VMT decreases with neighborhood density and with regional status. For example, even if living in a lower-density neighborhood in either case, the average household's annual VMT would be 10.7% lower (17,348) in a larger MSA with rail than in a smaller MSA (19,424). If living in a higher-density neighborhood in either case, the difference would be 8.2% (17,348 versus 18,907). The density of the residential neighborhood is also important: if living in a higher-density neighborhood in a smaller MSA, the average household's VMT would be 2.7% smaller (18,907) than if it lived in a lower-density neighborhood in the same region (19,424). Taken together, these results point to the roles played by both local density and regional status: even lower-density living can be associated with lower VMT for households that live in larger MSAs (especially those with rail), and even smaller regions can have lower VMT if residential neighborhoods are denser.

6.3 Accounting for the Impact of Residential Self-Selection

When analyzing the effects of land use characteristics on vehicle ownership and travel behavior, caution should be used in order not to overestimate the impact of the land use features on individual behaviors. This could happen if households self-select into neighborhoods that support their residential and mobility preferences: the observed differences in travel patterns would be attributed to differences in land use features, while (at least a part of) these differences may be actually associated with differences in individual/household preferences.

For example, individuals might choose to live in high density settings with varied land uses because they seek to drive less and they enjoy these types of urban settings. If this is true, they do not adopt these travel patterns entirely as a direct effect of the built environment, but also as a consequence of their personal attitudes and preferences. This residential self-selection effect may significantly reduce the effects of policies designed to reduce the use of private vehicles and incentivize alternative transportation modes. Mokhtarian and van Herick (2016) summarize the findings from several studies that applied propensity score and sample selection approaches to correct for residential self-selection. They conclude that the impact of built environment characteristics on travel behavior is often overestimated when residential self-selection is not controlled for: depending on the specific context, the true impact of the built environment can be as much as two-thirds lower than its estimate if residential self-selection is not controlled for.

Addressing the residential self-selection (RSS) issue is a key area of current research. A number of approaches to accounting for residential self-selection have been identified in the literature (Mokhtarian and Cao, 2008), each with advantages and disadvantages. Incorporating attitudes into an equation for vehicle ownership (or VMT) that already contains land use variables is one approach – referred to as "statistical control" – which is reasonably easy to understand and relatively effective. Again, however, we face the problem that the NHTS (although it contains a number of pertinent land use variables) does not contain data on many attitudes relevant to this particular issue. Conversely, the attitudinal datasets contain numerous appropriate attitude variables, but (1) for the most part, they include measures of vehicle ownership but not VMT, and (2) their land use measures are rather diverse and not always present for every case.

Another approach to accounting for residential self-selection is taken by Salon (2015), who modeled VMT as a function of land use and socioeconomic variables, through the joint estimation of residential location type choice (selection) and location-type-specific VMT models in a sample-selection approach to control for residential self-selection. By incorporating a correction term into the VMT model for each location category (the six clusters, in our study), this approach controls for the endogeneity bias caused by the omission of attitudes that influence both residential choice and travel behavior. In this study, we use a similar approach to control for self-selection. Because we are not aware of a study that has employed a multinomial selection model with a multinomial outcome model, we apply this approach only to VMT (which involves a regression outcome model), and not to vehicle ownership (which involves a multinomial outcome model).

To do this, we estimate a joint sample-selection model that includes a multinomial logit model of residential location, and six VMT log-linear models (each one for each neighborhood type) which include the sample selection bias correction term that accounts for the likelihood of a given household to live in such a neighborhood type. Several types of selection bias corrections have been proposed in the literature (Bourguignon et al., 2007). In the estimation of the sample-selection model for this project, we use the Lee correction term.

Table 6.5 summarizes the results of the multinomial-logit sample-selection model of residential location, which models the likelihood of a household to live in each of the six neighborhood type clusters that have been defined for this project, which was estimated with the NHTS dataset and using household size and composition and other sociodemographic characteristics as explanatory variables. In the estimation of this model we did not limit ourselves to the previous specifications, but found the best specification for this new model formulation (which brought new variables, such as education, into the model). The goodness of fit of this sample-selection model is rather low, which is not surprising because residential location is a complex decision that is affected by numerous individual- and household-specific variables that are unavailable in the NHTS dataset.

Table 6.6 presents the results of the best log-linear VMT model adjusted with the Lee sample-selection bias correction term to account for residential self-selection. While the results from this model are rather similar to those presented in Table 6.3 in significance and sign of the coefficients, the magnitudes of the coefficients are often somewhat different from the previous results. Thus, this model provides a more appropriate measure of the true impacts of these variables on household VMT, after accounting for households' residential self-selection into a specific neighborhood type. However, the bias correction term is sizable and strongly significant only for Cluster 2, which is unexpected and probably due to the low goodness of fit of the sample selection model.

Several conclusions can be drawn about the effect of the different explanatory variables from the analysis of the estimated coefficients for the ln(VMT+1) model for the six clusters. For example, as income increases for households in both high- and low-density neighborhoods, they tend to have higher VMT. The (positive) impact on VMT of increasing the number of driving workers is stronger for households in smaller towns than for those living in larger regions with rail, probably due to the presence of alternate modes of transportation in large cities and the greater dependence on cars in smaller towns. As the number of children under age 16 increases, the increase in VMT is roughly the same across all lower-density neighborhood clusters. However, for higher-density neighborhoods, the impact of children on VMT is larger for those households who live in smaller regions compared to those living in larger regions with rail transit. Again, greater dependence on cars is the likely reason for this phenomenon.

Table 6.5: Multinomial Logit Sample-selection Model of Residential Location (NHTS Dataset, Unweighted N = 121,839)

	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6
Income level 1 (less than $25,000)	*(base)*	**-0.891**	**-1.608**	**0.258**	**-0.624**	**-1.238**
		(0.029)	(0.042)	(0.066)	(0.053)	(0.055)
Income level 2 ($25,000 to 49,999)	*(base)*	**-0.668**	**-1.218**	**0.201***	**-0.416**	**-0.922**
		(0.024)	(0.032)	(0.058)	(0.044)	(0.044)
Income level 3 ($50,000 to 74,999)	*(base)*	**-0.545**	**-0.931**	**0.147***	**-0.424**	**-0.748**
		(0.024)	(0.031)	(0.058)	(0.045)	(0.045)
Income level 4 ($75,000 to 99,999)	*(base)*	**-0.374**	**-0.689**	0.001	**-0.413**	**-0.504**
		(0.025)	(0.032)	(0.065)	(0.050)	(0.047)
Driving workers (#)	*(base)*	**-0.139**	**-0.145**	**-0.281**	**-0.428**	**-0.396**
		(0.012)	(0.017)	(0.028)	(0.024)	(0.024)
Driving non-workers (#)	*(base)*	**-0.054**	**-0.050***	**-0.376**	**-0.459**	**-0.358**
		(0.013)	(0.018)	(0.030)	(0.026)	(0.026)
Non-driving workers (#)	*(base)*	-0.034	**0.223***	**0.367**	**0.381**	**0.526**
		(0.063)	(0.080)	(0.102)	(0.086)	(0.081)
Non-driving non-workers (#)	*(base)*	**0.100**	**0.275**	**0.101***	**0.149**	**0.371**
		(0.024)	(0.032)	(0.047)	(0.039)	(0.037)
Children (#)	*(base)*	**0.026***	**0.006***	**-0.108**	**-0.179**	**-0.162**
		(0.011)	(0.014)	(0.027)	(0.023)	(0.022)
Hispanic ethnicity (DV)	*(base)*	**-0.585**	**-0.980**	**-0.949**	**-1.396**	**-1.828**
		(0.036)	(0.044)	(0.063)	(0.049)	(0.046)
Asian (DV)	*(base)*	**0.979**	**1.692**	**1.083**	**1.473**	**2.267**
		(0.067)	(0.070)	(0.121)	(0.092)	(0.076)
Black (DV)	*(base)*	**0.451**	**0.375**	**0.198***	**0.725**	**1.029**
		(0.034)	(0.049)	(0.071)	(0.052)	(0.051)
Other race (DV)	*(base)*	**0.307**	**0.636**	**-0.173***	0.160	**1.014**
		(0.082)	(0.098)	(0.188)	(0.136)	(0.102)
Education Level 1	*(base)*	**-0.503**	**-0.838**	**-0.942**	**-1.090**	**-1.036**
		(0.045)	(0.075)	(0.087)	(0.081)	(0.082)
Education Level 2	*(base)*	**-0.306**	**-0.408**	**-0.680**	**-0.7000**	**-0.731**
		(0.024)	(0.034)	(0.051)	(0.045)	(0.047)
Education Level 3	*(base)*	**-0.051***	**-0.189**	**-0.460**	**-0.293**	**-0.370**
		(0.021)	(0.028)	(0.045)	(0.038)	(0.039)
Education Level 4	*(base)*	**0.160**	-0.030	**-0.240**	-0.011	**-0.085***
		(0.020)	(0.027)	(0.045)	(0.037)	(0.037)

House Tenure (Rent)	*(base)*	0.027 (0.027)	**0.218** (0.034)	**0.975** (0.041)	**1.131** (0.035)	**1.360** (0.036)
Constant	*(base)*	**1.237** (0.080)	**1.543** (0.098)	-0.046 (0.148)	**1.781** (0.116)	**2.650** (0.110)
Log likelihood final model:		-162140.1				
ρ^2 (equally-likely base):		0.0404				
Adjusted ρ^2 (EL base):		0.0399				

*Note: coefficients are in bold and standard errors are reported in parentheses. All estimated coefficients in bold are significant at least at the 0.1% level, unless otherwise noted. Coefficients marked with * are significant at the 5% level.*

Table 6.6: Best Model of ln(VMT+1) Adjusted with Sample-selection Bias Correction Term to Account for Residential Self-selection (NHTS Dataset, Unweighted N = 121,839)

	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6
Urban Density	-0.33554	-0.26653	-0.24457	-0.11781	-0.04326	-0.12037
Driving Limitation (DV)	-0.10569	-0.10035	-0.1114	-0.10372	-0.11112	-0.15570
Income Level 1 (less than $25,000)	-0.57409	-0.46472	-0.533	-0.5871	-0.49491	-0.37327
Income Level 2 ($25,000 to 49,999)	-0.30989	-0.25706	-0.36189	-0.32268	-0.31525	-0.29635
Income Level 3 ($50,000 to 74,999)	-0.10847	-0.06122	-0.10387	-0.13522	-0.13338	-0.09736
Income Level 4 ($75,000 to 99,999)	-0.07432	-0.04088	-0.02338	-0.06322	-0.04803	-0.04884
Number of Driving Workers	0.667858	0.654187	0.602013	0.679301	0.687166	0.650706
Number of Driving Non-Workers	0.454841	0.41047	0.387835	0.430067	0.455984	0.426807
Number of Non-Driving Workers	0.141517	0.114355	0.156373	0.170441	0.146829	0.011769
Number of Non-Driving Non-Workers	0.09249	0.045565	0.028254	0.063052	0.13875	0.085023
Number of Children Under Age 16	0.094186	0.081844	0.105361	0.115835	0.121571	0.123211
Cluster-Specific Sample-Selection Bias Correction Term (Lee's Correction Method)	0.004635	0.333847	0.067794	-0.06243	-0.03064	0.073853
Constant	8.640881	9.064952	8.84417	8.613353	8.60339	8.830163
R^2	0.3649	0.3830	0.3894	0.3248	0.3255	0.3091
Adjusted R^2	0.3648	0.3827	0.3889	0.3231	0.3243	0.3079
σ^2	0.615367	0.834908	0.563253	0.790818	0.698101	0.763745
ρ	0.005909	0.365366	0.090332	-0.0702	-0.03668	0.084507

7. Conclusions

A reduction in auto ownership likely reduces greenhouse gas emissions as well as traffic congestion. Therefore, understanding the factors behind owning fewer vehicles is crucial for implementing effective policies that can improve public welfare and reduce the environmental externalities of transportation. This study helps improve our understanding of why certain households choose to own fewer vehicles and drive fewer miles than usual. The study highlights the incremental ability of different model specifications, involving the inclusions of specific groups of variables in the vehicle ownership category and VMT models, to improve the ability to correctly predict choices in the model that were estimated. However, several of these variables, notably personal attitudes, are often not accounted for in vehicle ownership and travel behavior studies, thus limiting the ability of planning organizations to properly model and interpret the reasons behind these choices.

Several conclusions of interest to planners and policy-makers can be drawn from this study. First, after accounting for household composition, income, and driving limitations, the effects of attitudes on vehicle ownership appear to be relatively modest, but they are far from negligible. Controlling for individual attitudes improves the prediction of households' vehicle ownership patterns: individuals who prefer (a) transit over driving, (b) biking and walking over driving, as well as (c) shops within walking distance of their homes, tend to own fewer vehicles than expected. At the other end of the spectrum, individuals who report that they like driving and like to live in more spacious homes with larger yards are more likely to live in higher-than-expected vehicle-owning households.

Overall, those who live in ZVO or LTE households (1) are more likely to have attitudes supportive of a lower carbon footprint; (2) tend to have more such attitudes in combination; and (3) tend to hold those attitudes more strongly, compared to the rest of their peers. This indicates the value of policies directed at influencing attitudes to be more favorable toward sustainable lifestyle choices, but also suggests that it may take the combined effect of several such attitudes to change behavior – only holding one such attitude but not others may not suffice.

Not surprisingly, households belonging to the lower vehicle ownership categories (ZVO and LTE) are found, on average, to have lower income, more often include individuals who have driving limitations, and are more likely to live in rental housing units located in higher-density neighborhoods. These findings confirm that, in the general population, most households that do not own any vehicles do so out of necessity, because they have either limitations on their ability to drive or low enough incomes to limit their ability to own vehicles. However, when focusing on the households with higher income and no driving limitations (i.e., who have more space for voluntary changes in vehicle ownership), no large income differences are observed across vehicle ownership categories. This indicates that, beyond a certain income threshold, vehicle ownership decisions are largely made out of choice, and affected by non-income variables such as residential location, individual attitudes and lifestyle preference.

In particular, relative to others, ZVO households with higher income and no driving limitations tend to be more diverse, have comparable incomes to the households in the other vehicle ownership categories, consist of smaller households with fewer children (i.e., have higher income per capita), more often live in rental units in very high density neighborhoods, and drive fewer miles thanks to the increased accessibility of more central locations. The average population density of the neighborhoods where higher-income ZVO households live is more than four times the population density of HTE households' neighborhoods.

This study also confirms the importance of residential location and of the relationships between the characteristics of land use and vehicle ownership and VMT. Specifically, the estimation results for the models that control for residential neighborhood characteristics highlight how the key relationships between vehicle ownership-related choices and VMT vary for households living in different types of areas. For instance, having more children in the household contributes to greater VMT, and the estimated coefficient for the children variable is roughly the same across all lower-density neighborhoods. However, for higher-density neighborhoods, the presence of children under 16 in the household is associated with a larger increase in VMT in smaller regions than in large regions with rail. The presence of alternative transportation options in densely populated large regions may explain this phenomenon.

Further, as income increases, households become more similar to the highest-income households in their propensity to own vehicles (or not). However, the convergence between wealthy and less-wealthy households happens from different directions depending on the interaction between regional status and residential neighborhood density. In lower-density neighborhoods, as regional status diminishes (from larger region with rail to larger region without rail to smaller region) the less-wealthy become similarly likely to the wealthy to own cars (mostly out of necessity); in higher-density neighborhoods, as regional status *increases* the wealthy become similarly likely to the less-wealthy to not own cars (mostly out of choice).

Overall, and consistent with expectations, households that live in higher-density neighborhoods are more likely to be in the ZVO or LTE categories, and have lower VMT. Both local density and regional status matter: even lower-density living can be associated with lower VMT if located in larger MSAs (especially those with rail), and even smaller regions can have lower VMT if residential neighborhoods are denser. Neighborhood density has a stronger relationship with VMT (i.e., larger reductions in VMT are associated with a given increase in density) in lower-density neighborhoods than in higher-density ones. This difference is more pronounced in smaller regions than in larger ones. Specifically among higher-density neighborhoods, however, the strongest relationships between density and VMT are found in large regions with rail: a given density increase is associated with larger reductions in VMT for households living in rail-served regions, after controlling for the impacts of other variables.[12]

The results from the study indicate that policies designed to expand higher-density neighborhoods (and provide high-quality public transit) are associated with lower vehicle ownership and VMT. Creating a more effective biking infrastructure, which makes walking or

[12] It is important to note that while the impact of urban density was controlled for through the direct inclusion of that variable in most models (even those segmenting on neighborhood type), the influence of public transit on travel behavior was only accounted for in this study in an indirect way, through separating large MSAs with rail from those without rail and from smaller areas, in the definition of regional status and thence in the estimation of the best models. Accordingly, it is not possible to quantify the relative contributions of density versus the presence of transit in influencing travel behavior toward greater sustainability. Indeed, even if the level of transit service were accurately quantified and included in the models, it would tend to be highly correlated with density and therefore it would still be difficult to distinguish their separate influences.

biking much more attractive than driving, while not explicitly studied as part of this research, is also expected to produce similar results. Future policies that aim to reduce vehicle ownership and VMT should focus on improving public transit and expanding higher-density neighborhoods/ regions. The latter should also allow for the creation of a more effective biking infrastructure, which makes walking or biking much more attractive than driving. The presence of more shops, services, and businesses in close proximity with one another is also associated with lower VMT for many households, although the magnitude of the VMT reduction associated with a change in density varies depending on the local conditions of the neighborhood and the land use characteristics of the city, as shown in the analyses of this project. In particular, this effect is found to be non-linear, and its magnitude depends on the initial density values.

The results from the study highlight the importance, for future surveys on household vehicle ownership and travel behavior, of collecting information on individuals' attitudes, as they improve the ability to explain (and interpret) the complex behaviors associated with vehicle ownership and use – particularly the motivations behind voluntary reductions in the household's carbon footprint, which are a key interest of this study. The inclusion of attitudes would also greatly improve the ability to control for the effects of residential self-selection in assessing the impacts of the built environment on vehicle ownership and VMT.

As found in this study, individual attitudes account for a modest, although not trivial (12.4%), improvement in the explanatory power of the vehicle ownership model, beyond what can be explained by household composition, employment, income, or driving limitations. In addition, they shed valuable light on the motivations for individuals' voluntary choices regarding vehicle ownership and mobility patterns. The results of the study suggest that an increase in the voluntary reduction of vehicle ownership could result both from modifying land use patterns as well as through promoting attitudinal shifts toward more environmentally-beneficial mode choices – preferably in combination.

The current study is also subject to some limitations. With respect to the NHTS data, although the entire sample was weighted on six key variables to represent the California population, the resulting sample is likely not fully representative with respect to a number of unobserved characteristics (such as climate, as well as attitudes and local cultures). With respect to the attitudinal dataset, the lack of fully comparable attitudes across the different datasets that were available required various assumptions and estimations to be made, which caused a certain amount of error in the measurement of attitudes for each case, and which could partly account for the relatively modest role they played in explaining observed choices in this study. The attitudinal dataset also offered no information on VMT, and resource limitations precluded the acquisition of residential land use characteristics for that sample. Furthermore, the attitudinal dataset was dominated by Northern California residents, and (similar to the situation with the NHTS data) even after weighting the sample to represent California as a whole on the basis of several demographic characteristics, it may not be representative in terms of pertinent unobserved characteristics. These limitations could be addressed in future extensions of the research, and/or in future studies, e.g. through the development of a comprehensive data collection with purposely-designed surveys that collect information on a wider set of attitudinal dimensions, as well as on residential location, and both vehicle ownership and VMT.

References

Anowar, S., N. Eluru, and L. F. Miranda-Moreno (2014) "Alternative Modeling Approaches Used for Examining Automobile Ownership: A Comprehensive Review." *Transport Reviews,* **34(4):** 441-473.

Ben-Akiva, M. E., and S. R. Lerman (1985) *Discrete Choice Analysis: Theory and Application to Travel Demand*. MIT Press: Cambridge, MA.

Bhat, C. R. and J. Y. Guo (2007) "A comprehensive analysis of built environment characteristics on household residential choice and auto ownership levels". *Transportation Research Part B*, **41:** 506-526.

Bhat, C. R. and N. Eluru (2009) "A copula-based approach to accommodate residential self-selection effects in travel behavior modeling". *Transportation Research Part B: Methodological*, **43:** 749-765.

Blumenberg, E. and G. Pierce (2012) "Automobile Ownership and Travel by the Poor: Evidence from the 2009 National Household Travel Survey." *Transportation Research Record: Journal of the Transportation Research Board*, **2320**: 28-36.

Boer, R., Y. Zheng, A. Overton, G. K. Ridgeway, and D. A. Cohen (2007) "Neighborhood design and walking trips in ten U.S. metropolitan areas". *American Journal of Preventive Medicine*, **32:** 298-304.

Bourguignon, F., M. Fournier, and M. Gurgand (2007) "Selection bias corrections based on the multinomial logit model: Monte Carlo comparisons". *Journal of Economic Surveys*, **21(1)**: 174-205.

Cao, X., P. L. Mokhtarian, and S. L. Handy (2007) "Cross-sectional and Quasi-panel Explorations of the Connection between the Built Environment and Auto Ownership", *Environment & Planning A*, **39(4):** 830-847.

Cao, X., P. L. Mokhtarian, and S. L. Handy (2009) "Examining the Impacts of Residential Self-selection on Travel Behavior: A Focus on Empirical Findings". *Transport Reviews,* **29(3)**: 359-395.

Choo, S. and P. L. Mokhtarian (2004) "What Type of Vehicle do People Drive? The Role of Attitude and Lifestyle in Influencing Vehicle Type Choice". *Transportation Research A,* **38(3):** 201-222.

Circella, G., K. Tiedeman, S. Handy, F. Alemi, and P. L. Mokhtarian (2016) "What Affects U.S. Passenger Travel? Current Trends and Future Perspectives," White Paper from the National Center for Sustainable Transportation. University of California, Davis, February 2016, available online at https://ncst.ucdavis.edu/white-paper/ucd-ct-to-10-wp2-1a/ (last accessed on March 13, 2017).

Delbosc, A. and G. Currie (2012) "Choice and Disadvantage in Low-car Ownership Households." *Transport Policy*, **23:** 8-14.

De Jong, G., J. Fox, A. Daly, M. Pieters, and R. Smit (2004) "Comparison of car ownership models." *Transport Reviews*, **24(4):** 379-408.

De Waal, T., J. Pannekoek, and S. Scholtus (2011) *Handbook of Statistical Data Editing and Imputation*. John Wiley & Sons: Hoboken, NJ.

Federal Highway Administration (2011) "2009 National Household Travel Survey – User's guide", available online at http://nhts.ornl.gov/2009/pub/UsersGuideV2.pdf (last accessed on March 13, 2017).

Feng, Y., D. Fullerton, and L. Gan (2005) "Vehicle Choices, Miles Driven, and Pollution Policies," *NBER Working Paper 11553*.

Gao, S., P. L. Mokhtarian, and R. A. Johnston (2008) "Exploring the Connections among Job Accessibility, Employment, Income, and Automobile Ownership Using Structural Equation Modeling". *Annals of Regional Science,* **42(2):** 341-356.

Handy, S., X. Cao, and P. L. Mokhtarian (2005) "Correlation or Causality between the Built Environment and Travel Behavior? Evidence from Northern California." *Transportation Research Part D: Transport and Environment,* **10(6)**: 427-444.

Handy, S. and K. Clifton (2001) "Local Shopping as a Strategy for Reducing Automobile Travel." *Transportation*, **28(4):** 317-346.

Heffner, R. R., K. S. Kurani, and T. S. Turrentine (2007) "Symbolism in California's Early Market for Hybrid Electric Vehicles." *Transportation Research Part D: Transport and Environment,* **12(6):** 396-413.

Hensher, D. A., J. M. Rose, and W. H. Greene (2005) *Applied Choice Analysis: A Primer.* Cambridge, UK: Cambridge University Press.

Hotz, J. and R. Miller (1992) “Conditional Choice Probabilities and the Estimation of Dynamic Models”. *Review of Economics and Statistics*, **60(3):** 497-529.

Kalton, G. and I. Flores-Cervantes (2003). “Weighting methods”. *Journal of Official Statistics*, **19(2):** 81-97.

Mokhtarian, P. L. and X. Cao (2008) “Examining the Impacts of Residential Self-selection on Travel Behavior: A Focus on Methodologies”. *Transportation Research Part B*, **42(3):** 204-228.

Mokhtarian, P. L., I. Salomon, and L. S. Redmond (2001) “Understanding the Demand for Travel: It's Not Purely ‘Derived’.” *Innovation: The European Journal of Social Science Research*, **14(4)**: 355-380.

Mokhtarian, P. L., and D. van Herick (2016) “Quantifying residential self-selection effects: A review of methods and findings from applications of propensity score and sample selection approaches”. *Journal of Transport and Land Use*, **9(1)**: 9-28.

Neufeld, A. J. and P. L. Mokhtarian (2012) “A Survey of Multitasking by Northern California Commuters: Description of the Data Collection Process.” Institute of Transportation Studies, University of California, Davis, Research Report UCD-ITS-RR-12-32. Available online at http://www.its.ucdavis.edu/?page_id=10063&pub_id=1802 (last accessed on March 13, 2017).

Piatkowski, D. P., K. J. Krizek and S. Handy (2015) “Accounting of the Short Term Substitution Effects of Walking and Cycling in Sustainable Transportation.” *Travel Behaviour and Society*, **2(1)**: 32-41.

Potoglou, D. and Y. Susilo (2008) “Comparison of Vehicle-Ownership Models”. *Transportation Research Record*, **2076**: 97-105.

Salon, D. (2015) "Heterogeneity in the Relationship between the Built Environment and Driving: Focus on Neighborhood Type and Travel Purpose". *Research in Transportation Economics*, **52:** 34-45.

Salon, D., M. G. Boarnet, S. Handy, S. Spears, and Gil Tal (2012) “How Do Local Actions Affect VMT? A Critical Review of the Empirical Evidence.” *Transportation Research Part D: Transport and Environment*, **17(7)**: 495-508.

Schwanen, T., P. L. Mokhtarian (2007) “The Role of Attitudes toward Travel and Land Use in Residential Location Behavior: Some Empirical Evidence from the San Francisco Bay Area”. *Housing Policy Debate*, **18(1)**: 171-207.

Schwanen, T., P. L. Mokhtarian (2005) "What if You Live in the Wrong Neighborhood? The Impact of Residential Neighborhood Type Dissonance on Distance Traveled". *Transportation Research D*, **10:** 127-151.

Sivak, M. (2015) "Has Motorization in the US Peaked? Part 7 Update through 2013". Transportation Research Institute, University of Michigan, Ann Arbor.

Williams, T. A., H. Aamidala, J.D. Borowiec and B.A. Glover (2016) "Methodologies Used to Estimate and Forecast Vehicle Miles Traveled (VMT): Final Report". Texas A&M Transportation Institute, July 2016, available online at https://tti.tamu.edu/publications/catalog/record/?id=43634 (last accessed on March 13, 2017).

Ye, L., P. L. Mokhtarian, and G. Circella (2012) "Commuter Impacts and Behavior Changes during a Temporary Freeway Closure: the 'Fix I-5' Project in Sacramento, California." *Transportation Planning and Technology*, **35(3)**: 341-371.

Zhou, B. and K. Kockelman (2008) "Self-selection in Home Choice: Use of Treatment Effects in Evaluating the Relationship between the Built Environment and Travel Behavior". *Transportation Research Record*, **2077**: 54-61.